APPLICATIONS OF NEXT-GENERATION SEQUENCING TECHNOLOGIES AND ADVANCED ANALYTICAL METHODS TO UNDERSTAND GENE BY ENVIRONMENT INTERACTIONS

BY
KELSEY CAETANO-ANOLLÉS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Animal Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Sandra Rodriguez-Zas, Chair, Director
Professor Maria Villamil
Professor Matthew Wheeler
Assistant Professor Andrew Steelman

ABSTRACT

Addiction to activity and to psychoactive substances share psychologically and physiological characteristics. A transcriptome study compared a mouse line selected for high voluntary activity to a control line in environments restricting or enabling a rewarding activity. Results offered insights into genetic factors behind activity dependence and provided a model for understanding addiction and other reward-dependent behaviors.

Most genes differentially expressed between activity genotypes were only moderately differentially expressed between activity environments, suggesting that environmental effects were not confounded with activity genotype effects. Adora2a had a significant genotype-by-environment interaction effect evidenced by over-expression in the activity genotype relative to control in high activity environment and under-expression in the low activity environment. Our findings of differentially expressed genes related to dopaminergic transmission between activity genotypes support the association between these genes and activity. A central theme from the functional analysis of activity genotype-environment was neuron morphogenesis. Gene network analysis identified connected genes that exhibited similar (e.g. Lhx9) and opposite (e.g. Nrgn) expression patterns between activity genotypes across reward availability environment.

Our findings suggest that some transcriptomic changes in mice selected for high voluntary activity are shared with other reward-dependent behaviors. Results from this study support that high voluntary activity selection lines in mice are a helpful model to understand molecular mechanisms behind addition. Also, identification of genes and biological processes associated with both high voluntary activity and the pleasurable neurological response to physical activity may allow for the development of drugs which make it more pleasurable for people to exercise or less pleasurable to be sedentary as a treatment approach for overweight/obesity as well as a way to improve health overall in the general population.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

Introduction

Exploring gene expression is necessary for understanding the functioning of an organism at system level. In the past years, advances in next-generation sequencing (NGS) technologies such as RNA-sequencing (RNA-seq) have provided crucial tools to study changes in transcriptome abundance. My research will illustrate how these advances help understand behavior, gene-environment and genotype-phenotype interactions in model animal systems. The present chapter presents an overview of the main topics of my thesis.

In chapter II, I used data from an RNA-seq experiment to uncover interaction and main effects of Myostatin genotype and exercise on muscle transcriptome. Levels of myostatin expression and physical activity have both been associated with transcriptome dysregulation and skeletal muscle hypertrophy. The transcriptome of triceps brachii muscles from male mice (strain C57/BL6) corresponding to two genotypes (wild-type and myostatin-reduced) under two conditions (high and low physical activity) was characterized using RNA-seq. In chapter III, I explored differential gene and transcript abundance to characterize the cerebellum of mice demonstrating a high desire for physical activity relative to control that are restricted from or allowed to exercise. I am using a mouse model of dependency to study the transcriptome of reward-dependent behaviors.

Literature Review

The environment influences the expression of genes and molds how the molecular machinery of an organism and corresponding holobiont (host and symbiogenic microbiota) reacts to external and internal cues. Simple abiotic external stimuli such as temperature, light and chemicals, or biotic factors such as bacteria, archaea, fungi, and viruses that cause disease or are beneficial for the well-being of interacting organism, are among many factors that readily change patterns of gene expression as the organism

adapts to environmental change. The environment also triggers more durable changes that are hardwired into the organismal make up. These changes provide internal cues that are, for example, responsible for metabolic effects on disease (Hunter, 2005), hormonal control of sex and development (Pieau & Dorizzi, 2004), obesity-causing misregulation (Duhl et al., 1994), or regulatory control of social behavior (Zayed & Robinson, 2012). The advent of comparative genomics revealed that the number of genes of an organism is surprisingly invariant along vast organismal transects and that the differences in organismal complexity arise from elaborate mechanisms of regulation that impact the expression of their genes more than their genetic makeup (Levine & Tijan, 2003). Genome-wide studies have also shown that seemingly small changes in gene regulation result in large changes of morphology and physiology and that gene-environment regulation can cause extensive transcriptional rewiring and major modulation of transcript abundance (Tuch et al., 2008). Consequently, changes in gene expression patterns are primarily regulatory, express mostly at transcript abundance level, and respond appropriately to the generally highly dynamic interaction that exists between organisms and their environments. These changes affect for the most part the vast networks of protein-protein interactions that are responsible for the phenotype, which in humans are in the millions (Stumpf et al., 2008). Exploring gene expression is therefore necessary for understanding the functioning of an organism at system level. In the past years, advances in next-generation sequencing (NGS) technologies such as RNA-sequencing (RNA-seq) (Mortazavi et al., 2008) provided crucial tools to study changes in transcriptome abundance that are triggered by the many external and internal cues. This has changed the landscape of biological investigation, prompting a focus on bioinformatics, statistical genomics and system biology. The present research will illustrate how these advances help understand gene-environment interactions in model animal systems.

Gene Expression Microarray platforms

Microarrays have long been the most popular option for profiling global gene expression. Oligonucleotide ‘probes’ matching gene complements are synthesized and attached to specific positions of a microarray plate. Since each physical address of the

plate holds a specific oligonucleotide, addresses that "light up" indicate sets of genes present in DNA complementary to mRNA matching specific probes. While the use of microarrays is quick, cheap, and user-friendly, there is a substantial number of drawbacks to using this approach. These limitations have made finding an alternative method crucial. Firstly, hybridization methods such as those used in microarray profiling are limited by their reliance on existing genomic information. Microarrays must be built with in advance using known sequence information, meaning that the microarray resource becomes out-of-date as new genomic data is discovered and released. Additionally, this does not allow for discovery of new genes, since exploration is limited by the probes of the microarray plate. Hybridization-methods have high background levels and a limited dynamic range of detection, which results in low accuracy.

Gene Expression RNA-seq platforms

RNA-seq matches samples that you have, so your sample produces the actual outcome instead of using an external device limiting genomic features to be discovered. RNA-seq data can be reanalyzed an unlimited number of times as new information becomes available and updated tools and packages are released (Wang et al., 2009). The main benefit of RNA-seq approaches is that they allow for approximation of the abundance value of transcript (exactly how many copies of that transcript are present in the sample or biological endpoint). Sensitivity of microarrays is very low compared to RNA-Seq, so while you can measure increases or decreases in hydridization signal, it is impossible to numerically quantify changes in abundance relative to the entire transcriptomic profile. Microarray methods do not provide quantitative assessment of mRNA abundance nor dissect isoforms (Wang et al., 2009).

Tuxedo Suite and Galaxy

There are multiple methods available for RNA-seq; selection of a specific method for a study depends on various considerations, including factors such as transcript type, complexity or transcript boundaries. Transcriptome data appropriately processed can be analyzed using bioinformatics tools and software such as the TUXEDO suite, a collection

of applications to analyze differential gene expression (Langmead et al., 2009; Langmead & Salzberg, 2012; Trapnell et al., 2010; Trapnell et al., 2009). These applications are available free through Galaxy, a web-based gene expression analysis server (Goecks et al., 2010), which we make extensive use in the studies we here report. Use of this suite follows a Bowtie/Tophat -> Cufflinks -> Cuffmerge -> Cuffdiff workflow. Read alignment is performed by using either the Bowtie or Tophat applications depending on type of experiment and species studied. Bowtie, which is typically used to align shorter DNA reads to the human genome, is fast and memory efficient but not ideal for large genomes (Trapnell et al., 2012). Tophat, a fast splice junction mapper for RNA-seq reads, uses Bowtie to align reads to mammalian-sized genomes between 1 and 10 Gb (billions of base pairs per haploid genome) and identifies splice junctions between exons. Transcripts are assembled with the Cufflinks package, which also estimates abundances and tests for differential expression and regulation in RNA-seq samples. Cufflinks assembles aligned RNA-seq reads into transcripts and calculates estimates for the relative abundances of individual transcripts, using the number of reads supporting each transcript and controlling for library preparation protocol bias. After merging transcripts using Cuffmerge, Cuffdiff is used to compare the relative abundances of transcripts in various experimental conditions (Trapnell et al., 2012).

RNA-seq data analysis process

RNA-seq experiments follow a six-step pipeline (Van Verk et al., 2013), which is summarized here:

Step 1: RNA-seq library preparation

Before beginning data analysis, it is important to assess the quality of RNA, as poor quality can undermine the success of an experiment. Techniques such as Agilent Lab-on-Chip assay produce an RNA integrity number (RIN) between 1 and 10. A score of higher than 8 is typically considered favorable (Van Verk et al., 2013).

Step 2: Generating RNA-seq reads using Illumina-based sequencing

To perform RNA-sequencing, mRNA is first converted into cDNA; these millions of DNA fragments (reads) become the RNA-library, which is subsequently sequenced to obtain the measure of relative abundance of each individual transcript. (Van Verk et al., 2013) Currently, the most commonly used platform for RNA-seq analysis is the Illumina HiSeq platform. Total RNA is first quantified using a Nanodrop spectrophotometer and quality-assessed with an assay kit. NGS sequencing libraries are generated from around one microgram of total RNA using an RNA Sample Prep Kit (Illumina). Briefly, RNA is purified using oligo-T attached magnetic beads. After purification, the total poly A+ RNA is fragmented into small pieces. In the first strand of cDNA with random primers, the segmented mRNA fragments are reverse transcribed. These fragments are purified with a PCR extraction kit, resolved with elution buffer, and linked to sequencing adapters. The resulting libraries were then paired-end sequenced with the Illumina HiSeq™ 2000 system. Finally, complete paired end sequences are obtained as individual fastq files (forward and reverse) from output images with the CASAVA V 1.8.2 base calling software with ASCII Q-score offset 33 (Wang et al., 2009).

Step 3: Assess read quality, Alignment of RNA-seq reads

After these steps are performed, data is sent to the bioinformatics lab for downstream analysis in FastQ format, which includes both genomic and quality information. Quality is given in the format of a "PHRED" score (Q), calculated as $p=10^{(-Q/10)}$ where p is the probability that the corresponding base call is incorrect (An Extensive Evaluation of Read Trimming Effects on Illumina NGS Data Analysis). Adaptor sequences and low quality bases with PHERD scores (Q) $\leq$ 20 are typically removed (Wang et al., 2009). Read quality should be assessed using a tool like FastQC (Patel & Jain, 2012). FastQC allows data to be imported in various formats to check for unusual qualities that might indicate low sequence quality. Information is provided about the distribution of sequence length, overrepresented sequences, per sequence quality scores, and per base sequence quality. Summary graphs and tables are provided and quality can be assessed by visual inspection. *(*Patel & Jain, 2012*)*

At this point, if low quality portions of the NGS reads are present, read trimming can be performed to increase the quality and reliability of analysis. This is typically done by simply "surgically removing" the low quality bases (sequences?) and leaving the longest possible string of high quality sequence. Other common read preprocessing methods include duplicate removal, filtering contaminant sequences, or removal of adapter sequences. (Del Fabbro et al., 2013).

Reads are then aligned to either a reference genome, transcriptome, or *de novo* assembly. In cases where a reference genome is not available, for example if a non-model organism is being studied with no genomic information available, a *de novo* assembly can be created- this method blindly reconstructs the genome/transcripome using software such as Trinity (Haas et al., 2013). Mapping to a reference genome allows for precise measurements of transcript abundance and the identification of splicing isoforms. (Del Fabbro et al., 2013). For alignment to the transcriptome, Bowtie and BWA are most commonly used (Van Verk et al., 2013). Alignment to the transcriptome takes longer than alignment to the genome since it is more computationally intensive. These two aligners are not "splice aware", meaning that they cannot map reads over exon/intron junction. For alignment to a reference genome, splice aligner TopHat can be used (Van Verk et al., 2013).

Step 4: Assembly and quantification of gene expression

Cufflinks, part of the Tuxedo suite, assembles the mapped reads (Trapnell et al., 2010). Under Tophat, Bowtie previously flagged where reads match a location. Using the reference genome as a map, Cufflinks can assemble the reads where there is enough information to confirm the location. Next, the number of reads that map to each gene or transcript isoform are counted in order to quantitatively interpret our RNA-seq results. This count is referred to as "abundance". The more fragments of a transcript are sequenced, the more abundant that transcript is. Cuffmerge can merge multiple Cufflinks transcripts into a single transcriptome annotation file, which is useful for comparing results of different experimental conditions.

Step 5: Normalization of RNA-seq count data

Since most studies compare gene expression between different experimental

conditions and across samples, count data between different samples needs to be normalized in order to perform an accurate comparison. Differences in sequence depth and transcript length cause between-sample bias and within-sample bias, respectively. The most common normalization method is RPKM/FPKM (reads or paired-end fragments per kilobase of exon model per million mapped reads). This method both normalizes total gene length and the within-sample number of reads mapped, and is done in Cufflinks.

Step 6: Differential gene expression

Cuffdiff helps identify differences in transcript expression by calculating expression levels in more than one condition. They are then tested for significant differences. The FPKM (normalized) values from two conditions are divided from eachother and the log is taken. Next, the mean divided by the variance of this distribution provides a test statistic, which is used to calculate a P-Value and Q-Value (FDR adjusted p-value). A list of differentially expressed genes can be outputted and sorted by total expression levels in order to identify the most significant differentially expressed genes between conditions.

Gene ontology analysis using D.A.V.I.D.

Gene expression analyses give a large amount of extremely valuable data for downstream analysis. One crucial next step after gene expression profiling is to convert the large tables of gene lists into information that is adequately processed to biologically interpret results. The enrichment of functional categories and pathways among differentially expressed genes is one approach that can be explored using the web service Database for Analysis, Validation, and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov). DAVID makes use of functional annotations derived from several genomic resources, including the Gene Ontology (GO) project and the Kyoto Encyclopedia of Genes and Genomes (KEGG). The GO project (Ashburner et al., 2000) is a community effort to define ontological terms associated with gene functions and annotate relevant biological information in a hierarchy, under three main GO functional categories, biological process (BP), molecular function (MF), and cellular components

(CC). These high level terms are root nodes with no common parent nodes, meaning that terms of the annotation hierarchy will originate from each of these three root terms and will produce directed acyclic graphs (DAG), in which child terms can be associated to multiple parents to establish functional relationships and associations. Figure 1.1 shows the hierarchical structure of GO terms. BP represents the biological objective to which the gene or its product contributes, MF refers to the biochemical activity of a gene product, and CC refers to the location in a cell where the gene product is active (Ashburner et al., 2000). Pathways were those defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa et al., 2008). Minimization of redundancy within BP, MF and KEGG terms was attained by using GO Functional Annotation Tool (FAT) categories, which comprise more general BP and MF terms while filtering broad and less informative terms (Dennis et al., 2003). Statistical significance of the individual FAT or KEGG categories was based on the EASE Scores (modified Fisher Exact) available in DAVID. Further minimization of redundancy between FAT and KEGG categories was achieved through the use of clusters of categories that share genes. The statistical significance of these clusters was assessed using the enrichment score that corresponds to the geometric mean of the EASE Scores of the functional categories in the cluster (Huang da et al., 2007). Evaluation of clusters of categories enabled the discovery of molecular processes that respond to variation in physical activity and myostatin genotype levels.

Technical considerations in RNA-seq analysis

Most reads will map to a single location in the genome or transcriptome. However, in some cases, reads may align with more than one location. These are known as “multi-reads”. This can be attributed to duplicated or overlapping genes, or repetitive regions (Van Verk et al., 2013). Read aligners handle multireads by either randomly assigning them to one of the mapped locations, or equally distributing fractions of a read to each matched location. This causes gene regions with multireads assigned to them to have a different per base coverage than others, which subsequently causes under or over-estimation of the gene’s expression. These multireads are typically flagged in output and can be discarded, however, this leads to loss of information and accuracy when

estimating gene expression. Recent strategies have been developed that minimize information and power loss by assigning multireads to matched genes probabilistically. However, more research is needed to develop more effective ways to handle multireads in RNA-seq experiments (Van Verk et al., 2013).

De novo transcriptome assembly

Unfortunately, in some cases, no official reference genome for the species being studied exists. In this situation, there are effectively two representative approaches of profiling of gene expression using RNA-seq. One approach is for reads to be mapped to a related species. For example, in the case of quail, the closest reference genome would be that of the chicken. This approach can only measure expression of conserved regions in both species and is therefore unsuitable for precisely profiling the transcriptome. The second approach is making consensus by assembly reads to each other. There is a variety of assembly software available which can map the reads into assembled consensus sets without a reference genome for measurement of transcriptome levels.

Rationale for the use of mouse as a model system

The mouse (*Mus musculus*) has been selected as a useful model system for genomic, genetic and developmental research on mammalian species. While other organisms, such as yeast (*Saccharomyces cerevisiae*), worm (*Caenorhabditis elegans*) and fly (*Drosophila melanogaster*), represent excellent model systems for many biological and biomedical studies, especially for the study of development, the biology of the mouse has close genetic and physiological similarities to that of humans. Mammals arose 50-65 million years ago and rapidly diversified (Springer et al. 2004). They share a number of important biological features that unify many mammalian orders. For that reason, the mouse genome can be easily manipulated and analyzed using advanced techniques of genomic research for the study of the skeletal, nervous, immune, cardiovascular and endocrine systems of mammals. This helps the exploration of the molecular and physiological basis of numerous diseases that affect human and many

other mammalian organisms of economic and biological importance, including cancer, cardiovascular (e.g. heart and arterial diseases), metabolic (e.g. diabetes, obesity), neuromuscular (e.g. Parkinson's, multiple sclerosis), and skeletal (e.g. osteoporosis) disorders. Important diseases such as Alzheimer's, cystic fibrosis, and other neurodegenerative diseases, while not impacting the mouse, can be induced in the mouse and studied using the mouse model. Mice also represent excellent models for neurobiology and the study of behavior, including research on anxiety, aggressive behavior, drug and alcohol addiction.

A number of important resources support a vast scientific community exploring mouse biology, including integrated genetic, genomic and biology resource repositories (e.g. Mouse Genome Database; Eppig et al. 2014), gene expression (e.g. GXD; Smith et al. 2014), and developmental and functional databases (e.g. Mouse Atlas Project; Hayamizu et al. 2014) databases. The Mouse Genome Informatics initiative (MGI; http://www.informatics.jax.org) and the The Jackson Laboratory (public research center established in 1929) preserves and distributes mutant, transgenic and other variant mice, other resources, and important information for the mouse community. Genetically engineered mice (knockout mice) are available that have one or more genes made inoperable through gene knockout techniques. The Jackson Laboratory provides also genomic DNA and maintains a Bacterial Artificial Chromosome (BAC) library for distribution. MGI together with the International Mouse Strain Resource (http://www.findmice.org//index.jsp) provides information related to mutant genetic makeup, cryopreserved gametes, embryos or embryonic stem cells. Alleles of genes in gene trap libraries are available for example from the International Genetrap Consortium (IGTC; http://www.genetrap.org).

Adult mice weight 30-40 grams, multiply quickly (they reproduce as often as every 9 weeks) and have a relatively low cost of maintenance. Mouse are selected and bred to produce offspring with desirable traits utilizing the classical tools of genetics. Mice have been also exposed to DNA damage (irradiation and chemical mutagenesis) to produce variants that can be used as models of genetic disease. More recently, scientists have utilized numerous genetic technologies to produce 'mouse models' for the study of targeted genes and a wide array of specific diseases. Specifically, the creation of

transgenic mice, in which a new gene is inserted into the germline of the mouse, and the use of homologous recombination that permit to "knock out" and “knock in” genes (replacement of existing genes with altered versions in vitro or in situ), have generated invaluable information for our understanding the genome. There are over 450 inbred strains of mice providing information about their genotype and phenotype, and thousands of spontaneous, irradiation and chemically-induced mutants and transgenic lines. The Jackson Laboratory alone distributes 2,700 different strains and stocks as breeding mice, frozen embryos or DNA samples.

The mouse has 20 chromosomes in its haploid genome (i.e. 40 total chromosomes), with most chromosomes being acrocentric (not metacentric like in humans) (Figure 1.2). The mouse genome has a genome size of ~2,800 millions of base pairs (Mb) and a GC content of 42%. The genome holds 48,807 genes organized in 279 scaffolds encoding 78,107 proteins (updated 2/9/2015; NCBI; http://www.ncbi.nlm.nih.gov/genome/). There are 4.16 million records of curated gene expression profiles related to the mouse that have been deposited in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geoprofiles/) and ~272,000 gene expression datasets deposited in the GEO datasets repository (http://www.ncbi.nlm.nih.gov/gds/) of the NCBI. This wealth of information benefits the analysis of mouse gene expression patterns for the study of the transcriptome and the molecular biology of the cell.

Project Summary

My research will utilize NGS technologies to achieve the following three goals, each salient to the fields of animal sciences and bioinformatics:

(1) Use massive parallel next-generation RNA-seq methods to characterize the complete transcriptome of triceps brachii muscles from myostatin-reduced and wild-type/myostatin-typical mice at different physical activity levels. This work has been completed and published in *PLoS ONE* (Caetano-Anolles et al., 2015)

Myostatin is a hormonal protein responsible for inhibition of muscle growth and proliferation. Animals with Myostatin deficiencies show extremely large muscles and exhibit increased numbers and size of muscle fibers (McPherron & Lee, 1997). This is often referred to as 'double muscling.' While a substantial amount of research has been invested on the impact of Myostatin on muscle development during embryogenesis, very few studies have explored its effect when Myostatin is depleted post-development. This understanding is crucial, as researchers race for prevention and treatment of muscle diseases. In particular, genetic treatments to reduce Myostatin or exon skipping have potential to increase muscle mass in adults with disorders that affect muscle tone.

Understanding the genetic factors behind the function of Myostatin is essential in this regard. Our study utilized a 2 x 2 factorial approach, with two treatment groups (encouraged to or restricted from exercising) and two mouse strains (Myostatin-knockout and wild-type mice). RNA-seq data was gathered from the triceps brachii muscles, processed in Galaxy, and gene expression profiles subjected to downstream analyses with DAVID and collection of relevant gene ontology (GO) information. We found differential expression of genes not previously linked with Myostatin, related genes such as Follistatin, or genes associated with skeletal muscle function. While these genes are generally connected to muscle contraction and regulation, their expression reveals that they may play an important role in how Myostatin affects muscle development. For example, the highly abundant CYP26B1 is a gene responsible for lipid metabolism and synthesis. The gene has been shown to signal aortic smooth muscle cells through all-trans-retinoic acid regulation during cardiovascular development (Elmabsout et al., 2012; Ocaya et al., 2010; Nelson, 1999). There is no demonstration that CYP26B1 regulates aspects of muscle function and development other than those of a cardiovascular nature. Another highly expressed gene, TPM3, is essential for regulation of striated and smooth muscle contraction (Morris et al., 1991). Again, TPM3 has not been established to regulate any kind of muscle growth. These results reveal that genes once thought to exclusively control the contraction and function of fully developed muscles may have important roles in muscle development during myogenesis.

(2) Explore genomic variants at the RNA-level, transcription variants, and differential abundance of RNA-seq data obtained from the cerebellum of mice demonstrating a high desire for physical activity, similar to that of individuals with exercise dependence, and then either restricted from or allowed to exercise. (In progress)

The similarities between the biological reaction of the body to exercise and addiction-forming substances has been well established, laying the foundation for the use of exercise as a model for reward-dependent behaviors (Arias-Carrión et al., 2010; Hausenblas & Downs, 2002). The dopaminergic mesocorticolimbic system of the brain has a primary role in many aspects of reward-motivated response, such as associative learning or reinforcement and incentive salience (Berridge & Robinson, 1998; Chiara et al., 1999; Koob, 1992). The dopaminergic system has been shown to activate in studies where animals are provided with food or sexual stimuli, addictive substances, or electrical and self-stimulation, even reacting to secondary reinforcing stimuli (Ivlieva, 2011). Exercise causes a synthesis and release of dopamine in the basal ganglia (Meussen et al., 1997), indicating that it activates the dopaminergic mesocorticolimbic brain system and the neural processes of physical activity overlaps with those of other rewards substances. Given that exercise dependence and addiction to psychoactive substances share these similar characteristics psychologically and biologically, research on mice with a desire for physical activity similar to that of individuals with exercise dependence can not only give us insight into genetic factors behind the characteristic but can provide a model for understanding drug addiction.

3) In chapter IV, with the goal of using a mouse model of dependency to study the transcriptome of reward-dependent behaviors, I plan to explore differential gene and transcript abundance to characterize the cerebellum of mice demonstrating a high desire for physical activity relative to control that are restricted from or allowed to exercise. I plan to perform network analysis on the results of Chapter 3 to visualize the interactions between genes and associated addiction-related pathways to provide insight into similarities between addiction to drugs of abuse and addiction to exercise, as well as provide a foundation for better understanding gene interactions behind the unique phenotype and behavioral traits that selected mice in this experiment display.

FIGURES

Figure 1.1 The hierarchical structure of Gene Ontology terms.

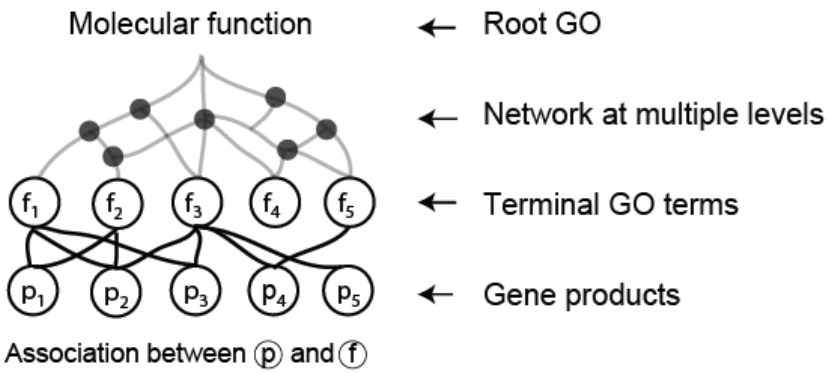
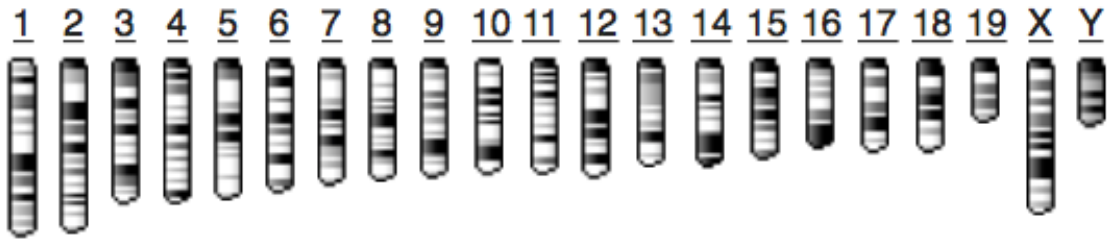


Figure 1.2 The 20 haploid set of the mouse genome (from NCBI Genome database). The X and Y chromosomes of the 20th set are shown.



CHAPTER 2: SYNERGISTIC AND ANTAGONISTIC INTERPLAY BETWEEN MYOSTATIN GENE EXPRESSION AND PHYSICAL ACTIVITY LEVELS ON GENE EXPRESSION PATTERNS IN TRICEPS BRACHII MUSCLES OF C57/BL6 MICE

Caetano-Anollés K, Mishra S, Rodriguez-Zas SL (2015) Synergistic and Antagonistic Interplay between Myostatin Gene Expression and Physical Activity Levels on Gene Expression Patterns in Triceps Brachii Muscles of C57/BL6 Mice. PLoS ONE 10(2): e0116828. doi:10.1371/journal.pone.0116828

Published manuscript available at:
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0116828

Abstract

Levels of myostatin expression and physical activity have both been associated with transcriptome dysregulation and skeletal muscle hypertrophy. The transcriptome of triceps brachii muscles from male mice (strain C57/BL6) corresponding to two genotypes (wild-type and myostatin-reduced) under two conditions (high and low physical activity) was characterized using RNA-seq. Synergistic and antagonistic interaction and ortholog modes of action of myostatin genotype and activity level on genes and gene pathways in this skeletal muscle were uncovered. The number of genes exhibiting significant (FDR-adjusted P-value < 0.005) activity-by-genotype interaction, genotype and activity effects were 1836, 238, and 399 genes respectively. The most common profiles were differential expression in (i) inactive myostatin-reduced relative to active and inactive wild-type, (ii) inactive myostatin-reduced and active wild-type, and (iii) inactive myostatin-reduced and inactive wild-type. Several remarkable genes and gene pathways were identified. The expression profile of nascent polypeptide-associated complex alpha subunit (Naca) supports a synergistic interaction between activity level and myostatin genotype, while Gremlin 2 (Grem2) displayed an antagonistic interaction pattern. Comparison between activity levels revealed expression changes in genes coding for structural proteins that are important for muscle function (including troponin, tropomyosin and myoglobin) and

genes important for fatty acid metabolism (some linked to diabetes and obesity, DNA-repair, stem cell renewal, and various forms of cancer). Conversely, comparison between genotype groups revealed changes in genes associated with G1-to-S-phase transition of the cell cycle of myoblasts and the expression of Grem2 proteins that modulate the cleavage of the myostatin propeptide. A number of myostatin-feedback regulated gene products that are primarily regulatory were uncovered, including microRNA impacting central functions and Piezo proteins that make cationic current-controlling mechanosensitive ion channels. These important findings extend hypotheses of myostatin and physical activity master regulation of genes and gene pathways, impacting medical practices and therapies associated with muscle atrophy in humans and companion animal species and genome-enabled selection practices applied to food-production animal species.

Introduction

Genetic and non-genetic conditions impact the molecular pathways and physiology of the skeletal muscle. The myostatin (Mstn) gene encodes a growth and differentiating factor and hormonal protein responsible for inhibition of muscle growth and proliferation in vertebrates. myostatin negatively regulates muscle fiber number during skeletal muscle development (McPherron & Lee, 1997; McPherron et al., 1997) and inhibits myogenic differentiation by reducing mRNA levels of the muscle regulatory factors (Rios et al., 2002; Matsakas et al., 2012). Conversely, myostatin deficiency caused by alterations at the DNA or RNA level is linked to increased numbers and size of muscle fibers, a phenomenon often referred to as ‘double muscling (Grobet et al., 1997). Likewise, myostatin-deficient Cre/loxP mice show hyperplasia (increased number of muscle fibers) and hypertrophy in the skeletal muscle (Heineke et al., 2010).

Physical activity influences muscle fiber in manners akin to the effect of myostatin deficiency (Bernardo et al., 2010). Contractile activity in muscle promotes changes in the synthesis and degradation of contractile and metabolic proteins that allow muscles to optimize, adapt, and endure activity (Kadi, 2000; Salmons & Vrbova, 1968; Cummins & Salmons, 1999). Activity causes an immediate sarcolemmal disruption that damages the cytoskeletal network. Muscle inflammation caused by physical activity is

accompanied by an increase of nitric oxide (NO) production, and skeletal muscle derived NO regulates contraction and metabolism as well as modulates muscle glucose uptake during activity (Kingwell, 2000).

While targeted genetic and non-genetic studies have associated skeletal muscle hypertrophy to dysregulation of the IGF1-Akt-mTOR and myostatin-Smad2/3 signaling pathways, muscle atrophy has been associated to dysregulation of the autophagic-lysosomal and proteasomal pathways (Schiaffino et al., 2013). Studies of the effect of myostatin deficiency using transgenic null myostatin (-/-) mice and quantitative real time PCR (qRT-PCR) identified over-expression of genes involved in myogenesis, protein degradation, extracellular matrix components and mitochondrial ATP synthesis (Zhao et al., 2009). Likewise, studies of the impact of physical activity on gene expression in the skeletal muscle using qRT-PCR demonstrated over-expression of myostatin and follistatin and under-expression of the myostatin receptor Activin IIB (ActRIIB) in murine limb muscles after acute physical activity (MacKenzie et al., 2013).

Few studies have evaluated the simultaneous effects of physical activity and myostatin genotype on the transcripts in the skeletal muscle of mice (Matsakas et al., 2012; Mosler et al., 2014). Results from one study demonstrated that the area of hypertrophic myofibres (extensor digitorum longus, gastrocnemius and rectus femoris muscles) in myostatin-depleted mice decreased towards wild-type levels meanwhile BCL2/Adenovirus E1B 19kDa Interacting Protein 3 (Bnip3), a key autophagy-related gene, was over-expressed in response to endurance exercise (Matsakas et al., 2012). Also, qRT-PCR profiling confirmed that activity increased the expression of the Uncoupling Protein 3 Mitochondrial, Proton Carrier (Ucp3), Carnitine Palmitoyltransferase 1A (Cpt1a), Pyruvate Dehydrogenase Kinase Isozyme 4 (Pdk4), and estrogen-related receptor-γ (Errγ) genes (Matsakas et al., 2012). Investigations focusing on the expression of a myostatin target gene, Mighty, using qRT-PCR suggested that acute resistance exercise decreased myostatin signaling in the skeletal muscle (soleus, plantaris and tibialis anterior muscles) of rats through the activation of the TGFβ inhibitor Notch. The activation of this inhibitor resulted in lower myostatin transcriptional activity that correlated with muscle hypertrophy (MacKenzie et al., 2013). A report of the effect of myostatin knockdown and exercise (swimming) in gastrocnemius muscle demonstrated

the over-expression of Pax-7 on knockdown mice without exercise relative to all other groups and over-expression of Myo-D on all knockdown mice irrespective of exercise level (Mosler et al., 2014). These previous studies have offered insights into the relationship between myostatin, physical activity and gene expression. However, studies that simultaneously consider genetic and non-genetic factors associated with the biological processes and molecular pathways in the skeletal muscle using high-throughput sequencing-based techniques could allow for a more comprehensive understanding of molecular networks of the skeletal muscle.

This study characterizes the complete transcriptome of triceps brachii muscles from C57/BL6 mice representing one of two genotype transcript levels (wild-type or myostatin typical and myostatin-reduced) and one of two physical activity levels (high and low) using massive parallel next-generation RNA sequencing. Synergistic, antagonistic and ortholog modes of action of the factors myostatin genotype and activity on genes and gene pathway profiles were investigated. This study is supported by: (a) mapping RNA sequencing reads to the mouse genome, identification of differentially expressed genes, and testing for differential expression among activity-genotype combination groups; (b) identification and interpretation of gene profiles revealing significant interaction between genotype and activity; (c) identification and interpretation of gene profiles revealing significant genotype (or activity) effect irrespective of activity (or genotype); and (d) functional analysis in support of the identification and interpretation of biological processes and pathways associated with genotype and activity levels. Our findings provide a basis to understand multifactorial gene regulation and dysregulation in triceps brachii and other skeletal muscles of mice.

Materials and Methods

Sample Collection

Profiling information stems from an experiment comparing the transcriptome of a skeletal muscles, triceps brachii muscle of adult (6 months of age) male C57/BL6 mice (Personal communication with Welle S., 2014). Difference in gene expression associated

with two factors were studied. The factor termed genotype encompasses two levels: wild-type mice exhibiting baseline expression of the myostatin gene and myostatin-reduced mice exhibiting lower expression of the myostatin gene. The factor termed activity encompasses two levels: inactive and active. Four physical activity-by-genotype combination groups of mice were compared (n=3/group): (1) active and myostatin-reduced, (2) inactive and wild-type (control genotype); (3) inactive and myostatin-reduced; and (4) active and wild-type. Prior to the trial, mice were housed in standard cages in groups of 2 or 3, given *ad libitum* access to food and water and kept in a 12-hour dark cycle. Myostatin-reduced mice were developed from C57/BL6 mice with floxed myostatin that was activated using Cre recombinase. At 4 months of age, all mice were fed chow with 0.025% tamoxifen content for 6 weeks to activate the Cre Recombinase enzyme and deplete myostatin only in mice with floxed myostatin genes. Tamoxifen is used to induce the Cre recombinase; it binds to the mouse estrogen receptor and disrupts its interactions with "chaperone proteins". Disruption of these interactions allows Cre-ER fusion proteins to enter the nucleus and perform recombination on the specific gene which has been floxed (Hayashi and McMahon, 2002). Myostatin was under-expressed (approximately 85%) in the myostatin-reduced mice. One week after the end of the tamoxifen feeding (approximately 6 months of age), mice in the active group were moved to be housed individually and given free access to running wheels during the last 12 weeks of the study. Physical activity was monitored and the sum over 1-hour periods was recorded. At the end of the wheel-running period, all mice were euthanized and samples taken from their triceps brachii muscles were frozen in melting isopentane and stored at −70°C. Muscles were sampled from myostatin-reduced and control mice that were matched for running behavior; this was done to ensure that the amount of physical activity performed was not a contributor to differences in gene expression between control and myostatin-reduced mice. Polyadenylated RNA was extracted as directed by Invitrogen using a Trizol reagent, converted to cDNA, and amplified with an Illumina TruSeq RNA library preparation kit following manufacturer instructions.

Triceps brachii muscle transcriptome was studied using Illumina Genome Analyzer IIx (Illumina, Inc. San Diego, CA) producing 65-base long single-end reads. Data processing was performed using CASAVA software. The 65-base sequence reads

were mapped to the mouse genome (mm9) using default settings and reads mapping to exons in the Refseq database (http://http://www.ncbi.nlm.nih.gov/refseq/) of transcripts were counted. Exon-level data was consolidated to gene-level read counts and summarized. Data normalization was performed so trimmed column means (excluding the highest and lowest 5th percentiles) were equal for all samples. In total, 13 protein-coding mtDNA transcripts were mapped to NC_005089.1, counted separately, and normalized to the trimmed mean of the non-mitochondrial transcripts. The transcriptome data, additional experimental details and preliminary analysis are available in the National Center for Biotechnology Information, Gene Expression Omnibus database, accession number GSE31839 (Edgar et al., 2002). Results from these preliminary analysis uncovered a main effect of activity on genes associated with oxidative energy metabolism and no interaction between activity and myostatin levels. Results from advanced modeling and functional analysis of the experiment are presented.

RNA-seq using Galaxy

The RNA-seq workflow generally involves five sequential steps: (i) Establishing an experimental design that surveys a number of biological endpoints used for hypothesis testing; (ii) Isolation and purification of RNA (generally mRNA, but sometime small RNA), (iii) Conversion of RNA into cDNA and addition of sequencing adapters; (iv) cDNA sequencing using one of the many NGS platforms; and (v) Assembly of the short-read sequences into transcription profiles and analysis. Here we focus on the assembly, statistical and downstream analysis of gene expression patterns (Figure 2.1).

There are multiple methods available for RNA-seq; selection of a specific method for a study depends on various considerations, including factors such as transcript type, complexity or transcript boundaries. Transcriptome data appropriately processed can be analyzed using bioinformatics tools and software such as the TUXEDO suite, a collection of applications to analyze differential gene expression (Langmead et al., 2009; Langmead & Salzberg, 2012; Trapnell et al., 2010; Trapnell et al., 2009). These applications are available free through Galaxy, a web-based gene expression analysis server (Goecks et al., 2010), which we make extensive use in the studies we here report. Use of this suite

follows a Bowtie/Tophat -> Cufflinks -> Cuffmerge -> Cuffdiff workflow. Read alignment is performed by using either the Bowtie or Tophat applications, depending on type of experiment and species studied. Bowtie, which is typically used to align shorter DNA reads to the human genome, is fast and memory efficient but not ideal for large genomes (Trapnell et al., 2012). Tophat, a fast splice junction mapper for RNA-seq reads, uses Bowtie to align reads to mammalian-sized genomes between 1 and 10 Gb (billions of base pairs per haploid genome) and identifies splice junctions between exons. Transcripts are assembled with the Cufflinks package, which also estimates abundances and tests for differential expression and regulation in RNA-seq samples. Cufflinks assembles aligned RNA-seq reads into transcripts and calculates estimates for the relative abundances of individual transcripts, using the number of reads supporting each transcript and controlling for library preparation protocol bias. After merging transcripts using Cuffmerge, Cuffdiff is used to compare the relative abundances of transcripts in various experimental conditions (Trapnell et al., 2012).

Data Analysis

The 65-base, single-end sequence reads from FastQ files were mapped to the mouse mm10 genome assembly accessed from the UCSC Genome Browser database (http://genome.ucsc.edu). Prior to mapping, FastqGroomer was used to convert file format to FastqSanger (Blankenberg et al., 2010) and FastQC was used for quality control of the reads (Patel & Jain, 2012). Using Fastq Quality Trimmer, 3' end positions that exhibited Phred quality values < 20 were removed. Data was normalized so trimmed transcript count means, excluding values in the upper and lower 5 percentiles, were the same for all samples. The 13 protein-coding mtDNA transcripts were mapped to `NC_005089.1 (*Mus musculus* mitochondrion, complete genome), counted separately, and normalized to the trimmed mean of non-mitochondrial transcript counts. Sequence reads were mapped using Tophat (v1.4.0), transcript isoforms were identified, quantified in number of fragments per kilobase of exon per million mapped reads (FPKM; Vivier et al., 2008), and differential transcript abundance was tested using Cufflinks routines including Cuffmerge, and Cuffdiff (v2.1.1) with default settings (Trapnell et al., 2012).

Differential expression was tested between activity-genotype groups, between activity groups and between genotype groups to determine the statistical significance of interaction and main effects on individual transcript isoforms (Le Bechec et al., 2011; Allison et al., 2002). False discovery rate (FDR) adjusted P-values were used to account for multiple test adjustment across transcripts (Benjamini & Hochberg, 1995; Storey, 2002; Storey et al., 2004). The vast majority of the differentially expressed genes were represented by one transcript and thus, results are discussed on a gene basis. The routine workflow was implemented in Galaxy (Goecks et al., 2010; Blankenberg et al., 2011; Giardine et al., 2005). Comparisons of lists of differentially expressed genes among contrasts between pairs of activity-genotype combinations were visualized with Venn diagrams, created using VENNY, an online interactive tool for comparing lists (Oliveros, 2007).

Gene ontology analysis using D.A.V.I.D.

Enrichment of functional categories and pathways among the differentially expressed genes was explored using the web service Database for Analysis, Validation, and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov). The Gene Ontology (GO) functional categories investigated included biological process (BP) and molecular function (MF), the biological objective to which the gene or its product contributes or the biochemical activity of a gene product, respectively (Ashburner et al., 2000), and pathways were those defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa et al., 2008). Minimization of redundancy within BP, MF and KEGG terms was attained by using GO Functional Annotation Tool (FAT) categories, which comprise more general BP and MF terms while filtering broad and less informative terms (Dennis et al., 2003). Statistical significance of the individual FAT or KEGG categories was based on the EASE Scores (modified Fisher Exact) available in DAVID. Further minimization of redundancy between FAT and KEGG categories was achieved through the use of clusters of categories that share genes. The statistical significance of these clusters was assessed using the enrichment score that corresponds to the geometric mean of the EASE Scores of the functional categories in the cluster (Huang da et al., 2007).

Evaluation of clusters of categories enabled the discovery of molecular processes that respond to variation in physical activity and myostatin genotype levels.

Identification of Synergistic and Antagonistic Gene Expression Patterns

Beyond the identification of differentially expressed genes exhibiting significant physical activity-by-genotype interaction, this study aimed at uncovering the synergistic or antagonistic interplay between these factors. Six pairwise contrasts were used to profile the expression patterns: active wild-type vs inactive myostatin-reduced [AW-IM], active myostatin-reduced vs active wild-type [AM-AW], active myostatin-reduced vs inactive myostatin-reduced [AM-IM], inactive wild-type vs active wild-type [IW-AW], inactive wild-type vs active myostatin-reduced [IW-AM], and inactive wild-type vs inactive myostatin-reduced [IW-IM]. Among the genes that exhibited significant (FDR-adjusted P-value < 0.001) activity-by-genotype interaction, alternative profiles of over- and under-expression or non-significant (raw P-value < 0.00005 or FDR-adjusted P-value < 0.05) differential expression in each of the six contrasts between pairs of activity-genotype combinations were identified. The concept of synergism and antagonism has been used previously in studies on expression and regulation (Lutter et al., 2010). While positively correlated gene expression patterns indicate synergism, anti-correlated or uncorrelated expression patterns indicate antagonism between genes. In this study, a synergistic interaction was detected when the effect of a particular combination of genotype and activity levels on gene expression was more than the sum of the effects of genotype and activity levels considered independently. In other words, synergistic effects were identified when the expression of a gene under a combination of genotype and activity levels was more extreme than the average expression under each level separately. Likewise, an antagonistic interaction was detected when the effect of a particular combination of genotype and activity levels on gene expression was less than the sum of the effects of genotype and activity levels considered independently. In other words, antagonistic effects were detected when the expression of a gene under a combination of genotype and activity levels was less extreme than the average expression under each level separately. Distinct profiles including a minimum of 50 genes were

further evaluated and functional categories enriched within cluster were investigated in DAVID.

Presentation of findings

Findings are presented and discussed in a sequence starting with genes exhibiting significant activity-by-genotype interaction, followed by genes exhibiting significant main effects of activity or genotype. Genes corresponding to transcripts exhibiting significant (FDR-adjusted P-value < 0.005; log2(fold change) > |1.3|) differential expression were identified in the Results section and their profile is discussed in the Discussion sections. Broader lists of genes reaching a lower significance threshold are presented in the supplementary materials (in File S1). Discussion focuses on gene expression patterns previously unreported in the context of the conditions studied and on patterns previously reported in similar or comparable studies.

Results

RNA-seq analysis and organization of findings

Considering that myostatin inhibition and physical activity are being explored as treatment options for muscle degeneration and other disorders, it is important to understand the impact of these factors at the gene co-regulation level. The RNA-seq profile analyses revealed changes in the transcriptome of a skeletal muscle, the triceps brachii muscle, between C57/BL6 wild-type and myostatin-reduced mice under two physical activity conditions. First, the quality and quantity of the sequence reads was evaluated across samples. The average size of the RNA-seq FastQ file was 1.3 G bases/sample. The average quality score Phred of the reads along the 65 positions across all samples was 30. The number of reads and quality scores along the reads were comparable across samples from all four activity-by-genotype groups. Likewise, the percentage of reads mapped to the mouse genome was similar across samples and was on

average 84.6% (17,486,782 of 20,675,801 total reads mapped). Of these, 5,013,631 (28.7%) had multiple alignments (12,183 had >20 alignments).

Activity-by-genotype interaction effect on gene expression in triceps brachii muscles

Overall, 1,836 genes exhibited significant (FDR-adjusted P-value < 0.005; maximum log2(fold change) > |1.3|) activity-by-genotype interaction. Table 2.1 lists genes exhibiting significant (FDR-adjusted P-value < 2XE-12) activity-by-genotype interaction effects due to space limitations. An extended list of differentially expressed genes with FDR-adjusted P-value < .01 is presented in Table S2.1 (supplementary table within published manuscript); in total, 84 genes were differentially expressed at this cutoff. The most extreme average fold change among genes exhibiting significant interaction was observed for the contrasts IM-AW, followed by IW-IM and AM-IM. This result indicates that the expression of genes in the inactive myostatin-reduced mice tended to be most different than the other three activity-genotype groups and this profile is a driving factor on the identification of genes expression exhibiting significant interaction. Conversely, on average the less extreme fold change among genes exhibiting significant interaction was observed in the contrast IW-AM. This finding reveals that these two conditions do not exhibit a synergistic effect among the genes presenting significant interaction.

Patterns of differential gene expression across activity-genotype contrasts

The number of differentially expressed genes (P-value < 0.05) for AW-IM, AM-AW, AM-IM, IW-AW, IW-AM, and IW-IM was 1,051, 86, 711, 119, 238 and 390, respectively. Several interpretations can be made from the progression of number of differentially expressed genes starting with the highest number in AM-IM followed by AM-IM, followed by IW-IM followed by AM-AW. Firstly, activity level was associated with more differentially expressed genes in myostatin-reduced mice than in the wild-type mice. Also, genotype was associated with more differentially expressed genes in active relative to inactive mice. Among all activity-genotype combination groups, inactive

myostatin-reduced mice exhibited the most number of differentially expressed genes relative to other activity-genotype combinations, including the highest number of over-expressed genes at FDR-adjusted P-value < 0.01**.** On the other hand, active wild-type mice exhibited the fewest number of differentially expressed genes relative to all other activity-genotype combinations, and these genes were overexpressed at FDR-adjusted P-value < 0.01.

Among active mice, the genotype difference was associated with the fewest number of differentially expressed genes among all pairwise contrasts. Likewise, among wild-type mice, activity level was associated with the second lowest number of differentially expressed genes among all pairwise contrasts. In contrast, changes in activity level elicited more differentially expressed genes in myostatin-reduced mice than in wild-type mice.

Figure 2.2 presents a Venn diagram illustrating shared differentially expressed genes between one set of three orthogonal contrasts including the IW baseline group. Of these genes, 33 genes were shared between all three contrasts including the IW baseline group (Table 2.2). These genes are highlighted because their expression in IW (inactive wild type) was significant different from all other three groups. These genes are of interest because either one or both conditions (genotype and activity) resulted in a departure from baseline conditions. The majority of the genes differentially expressed between IW-AM were shared with the contrasts IW-AW and IW-IM (130 out of 238). This same pattern was evident in the contrast of IW-AW to IW-AM and IW-IM (90 out of 119). In contrast, the majority of the genes differentially expressed in the contrast IW-IM were unique to this contrast (263).

Functional enrichment analysis of activity-genotype contrasts

Functional enrichment analysis was performed on genes exhibiting differential expression between pairs of activity-genotype combination groups (FDR-adjusted P-value < 0.01) for each of the six contrasts individually. Clusters of categories exhibiting enrichment scores > 3.0 (corresponding to average across categories within a cluster P-value < 0.01) were considered for discussion. These functional enrichment results can be

found in Tables S2.2 to S2.6 (supplementary tables within published manuscript). The contrasts AM-IM (Table S2.2) and IW-AW (Table S2.3) shared the cardiac muscle contraction (mmu04260) KEGG pathway, indicating that changes in activity level are associated with differential expression of genes linked to the muscle contraction network, regardless of genotype. The differentially expressed genes in these contrasts also exhibit enrichment of the tricarboxylic acid cycle TCA cycle (mmu00020) KEGG pathway and a number of GO MF terms linked to metabolism of cofactors and vitamins that are also linked to these metabolic pathways.

Notably, some functional categories enriched among the genes differentially expressed in the IW-IM contrast (Table S2.4) were also enriched among genes differentially expressed in the AM-AW (Table S2.5). Amid these, the BP categories of muscle cell differentiation (GO:0042692), various muscle development terms, and the KEGG pathways hypertrophic and dilated cardiomyopathy (mmu05410 and mmu05414, respectively). Also, the BP vasculature development (GO:0001944) was enriched among the genes differentially expressed in the AM-IM (Table S2.2) and IW-IM (Table S2.4) contrasts.

Genes exhibiting significant activity-by-genotype interaction

Overall, 2,074 genes exhibited a significant activity-by-genotype interaction association with expression (FDR-adjusted P-value < 0.01) The genes exhibiting the most significant interaction effect including the log2 fold change between pairs of activity-by-genotype groups and the overall FDR-adjusted P-value are presented in Table 2.1. A more extensive list of genes is provided in Table S2.1 (supplementary table within published manuscript). Several genes that exhibited a significant interaction also had one or more significant individual pairwise contrasts between activity-genotype combinations. Examples of these genes included: methyltransferase like 21E gene (Mettl21e); Tropomyosin 3 (Tpm3); Troponin T1 slow skeletal muscle (Tnnt1); nascent polypeptide-associated complex alpha polypeptide (Naca); dual specificity phosphatase 18 (Dusp18) and ATPase, Na+/K+ transporting, beta 2 polypeptide (Atp1b2).

Other genes that exhibited a significant interaction did not reach high significance in particular contrasts; however, the integration of consistent borderline significant contrasts resulted in a significant overall interaction effect (Table 2.1 and Table S2.1). Examples of these genes included: ATPase, Na+/K+ Transporting, Beta 2 Polypeptide (Atp1b2), p21 protein (Cdc42/Rac)-activated kinase 1 (PAK1), and Nascent Polypeptide-Associated Complex Alpha Subunit (Naca). The interaction pattern of PAK1 is characterized by highest expression in active myostatin-reduced, followed by inactive-wild type, followed by inactive myostatin-reduced mice.

The interaction pattern of cysteine and glycine-rich protein 3 (CSRP3) and Myozenin 2 (Myoz2) were characterized by highest expression in inactive myostatin-reduced mice relative to all other activity level-genotype groups. The consistent interaction patterns of Myosin, light polypeptide 2 (Myl2) and myosin, light polypeptide 3 (Myl3) can be summarized by over-expression in inactive myostatin-reduced mice relative to all other activity level-genotype groups. The pathways of these Myosin genes could result in this consistent profile. Myosin light chains (Myls) modulate muscle contraction and may be involved in myogenesis or muscle regeneration (Zhang et al., 2009; Rushforth et al., 1998; Andruchov et al., 2006; Timson, 2003).

Clusters of expression profiles among genes exhibiting significant activity-by-genotype interaction

Within the genes that exhibit significant interaction, various profiles of over- and under-expression were identified. However, the joint functional analysis of these groups may hinder or bias the discovery of enriched categories. Thus, discretizing the pairwise contrasts into over-, under- and non-differentially expressed identified clusters of genes exhibiting common expression profiles across the six contrasts. Profiles including more than 100 genes were subjected to functional enrichment analysis in DAVID. Table 2.3 lists the most common profiles of over- (denoted with a 1), under- (denoted with a -1) and not differentially expressed genes across the 6 pairwise contrasts considered and the number of genes in each profile cluster. The three most common profiles were: +100000 (183 genes), -100000 (146 genes), and -10-1000 (142 genes), where the first column in

the series of +1, -1 and 0 denotes the AW-IM contrast and the third column denotes the AM-IM contrast. For example, +100000 represents a pattern of significant positive differential expression in the AM-IM contrast and non-significant (P-value > 0.0001) differential expression in all other contrasts.

Tables 2.4, 2.5, and 2.6 list the clusters of functional categories enriched for profiles +100000 and -100000 (enrichment score > 3.0), and -10-1000 (enrichment score > 2.0). Two of the differential expression profiles across the pairwise contrasts included over- or under-expression in the two most extreme activity-genotype contrasts studied IW-AM (categories +100000 and -100000). Considering that genes from a process or pathway can be impacted in opposite manner by the same activity-genotype combination, the genes in the previous two profile clusters were combined and functional analysis of this augmented cluster was also undertaken.

Gene under-expression in active wild-type and active myostatin-reduced relative to inactive myostatin-reduced is a clear example of activity-by-genotype interaction because the combination of inactivity and myostatin-reduced genotypes was associated with higher gene expression than activity (regardless of genotype). Among genes sharing the first profile (under-expression in active wild-type and active myostatin-reduced relative to inactive myostatin-reduced) one highly enriched functional cluster (enrichment score = 16.28) was identified (Table 2.4). This cluster is comprised of 14 GO BP, MF, and KEGG pathway terms including the cardiac muscle pathway and pathways for inflammation-associated neurodegenerative conditions including Parkinson's disease, Alzheimer's disease, and Huntington's disease.

Among the profile cluster characterized by genes under- or over-expressed in the AW-IM contrast and not differentially expressed in all other contrasts (Table 2.5), five functional clusters presented enrichment scores > 2.0. Functional categories in the two clusters with highest scores are discussed (Table 2.5). The first functional cluster (enrichment score = 5.06) contained 9 BP terms. Oxidation reduction and oxidative phosphorylation were again found significant in this list along with multiple electron transport terms including electron transport chain, respiratory electron transport chain, ATP synthesis coupled electron transport, and mitochondrial ATP synthesis coupled electron transport. Additionally, this list includes generation of precursor metabolites and

energy, cellular respiration, and energy derivation by oxidation of organic compounds terms. The second functional cluster (enrichment score = 4.18) contained 15 GO BP, MF, and KEGG pathway terms. This cluster contained many of the same terms found significant in the previous contrast discussed, including oxidative phosphorylation, cardiac muscle contraction, and neurodegenerative diseases. A large number of electron transport chain, energy generation, and mitochondrial ATP synthesis terms appeared on this list as well, highlighting again the crucial electron-transfer role of mitochondrial-driven processes.

Among the profile characterized by genes under- or over-expressed in the AM-IM contrast and not differentially expressed in all other contrasts (Table 2.6), the functional cluster presenting an enrichment score > 2.0 is discussed. The first functional cluster (enrichment score = 2.43) contained 5 BP terms. These significant terms included: angiogenesis, blood vessel development, vasculature development, blood vessel morphogenesis, and tube development. These categories are consistent with the vascular development category previously identified.

Main effect of myostatin genotype on gene expression in triceps brachii muscles

Table 2.7 lists the genes differentially expressed (FDR-adjusted P-value < 0.005; log2(fold change) > |1.3|) between genotype groups excluding transcripts exhibiting significant interaction effects. An extended list of differentially expressed genes (FDR-adjusted P-value < 0.01) is presented in Table S2.7 (supplementary table within published manuscript). In total, 230 genes were differentially expressed between genotypes at FDR-adjusted P-value < 0.005. The most significant differentially expressed genes (Table 2.7) had positive log2(fold change) estimates indicating over-expression of those genes in the myostatin-reduced relative to wild-type mice.

Clusters of functional categories enriched (enrichment score > 3.0) among the genes differentially expressed between genotypes (raw P-value < 0.00005 comparable to FDR-adjusted P-value < 0.005) are listed in Table 2.8. An extended list of clusters (enrichment score > 1.5) including gene membership is presented in Table S2.8 (supplementary table).

Main effect of physical activity on gene expression in triceps brachii muscles

Table 2.9 lists the genes differentially expressed (FDR-adjusted P-value < 0.005; log2(fold change) > |1.3|) between triceps brachii muscle from active and inactive mice excluding genes exhibiting significant interaction effects. An extended list of differentially expressed genes between activity levels (FDR-adjusted P-value < 0.01) is presented in Table S2.9 (supplementary table). In total, 384 genes were differentially expressed at FDR-adjusted P-value < 0.005. The most significant differentially expressed genes (Table 2.9) had positive log2 (fold change) estimates indicating over-expression of those genes in active relative to inactive mice.

Clusters of functional categories enriched (enrichment score > 3.0) among the genes differentially expressed between activity levels (raw P-value < 0.00005 or FDR-adjusted P-value < 0.005) are listed in Table 2.10. An extended list of clusters enriched among genes exhibiting significant activity effect (including gene membership) is presented in Table S2.10 (enrichment score > 2.0).

Both genotype and activity level were associated with significant changes in gene expression, irrespective of the remainder factor indicating main effects. The more extreme fold change estimates observed in the genotype relative to the activity contrasts (based on the top differentially expressed genes) indicate that the genotypes considered in this study have a higher impact on gene expression than the physical activity levels evaluated (Tables 2.7 and 2.9).

DISCUSSION

Understanding activity-by-genotype interaction effect on gene expression in triceps brachii muscles

Common expression profiles. The patterns of differential gene expression across activity-genotype contrasts provided insight into activity-by-genotype interaction effect

on gene expression in triceps brachii muscles. The identification of 33 genes shared between one set of the three orthogonal contrast groups including the baseline group set, inactive wild-type mice (IW-AM, IW-AW, and IW-IM) suggests a specific role of these genes in inactive wild-type mice that are sensitive to changes in the activity-genotype condition (Table 2.2). Meanwhile, the large percentage of non-overlapping genes in the contrast of inactive wild-type relative to both active and inactive myostatin-reduced offers a glimpse to the distinct impact of activity level on gene expression within myostatin-reduced mice.

Consideration of the number of differentially expressed genes across pairwise contrasts alone uncovered insightful interaction patterns that are cornerstone for more complex patterns across contrasts. Among all pairwise contrasts, myostatin-related genotype differences within the active group (contrast AM-AW) were associated with the fewest number of differentially expressed genes (86 genes), suggesting that activity may be picking up and modulating or compensating some expression regulated by myostatin. The second lowest number of differentially expressed genes (119 genes) was related to differences in activity level within wild-type mice (contrast IW-AW), indicating that activity alone was associated with more limited changes in gene expression than those observed in the combination of activity and myostatin reduction. Finally, considering the fact that activity level elicited more differentially expressed genes in myostatin-reduced than in wild-type mice along with the finding that the greatest number of differentially expressed genes is in the AW-IM contrast confirms the hypothesized synergistic impact of physical activity and the silencing of myostatin on gene expression.

Functional enrichment within profiles. Functional enrichment analysis of the activity-genotype contrasts offered insights into the impact of the activity-genotype combinations on GO MF and BP categories and KEGG pathways. Differentially expressed genes in the AM-IM and the IW-AW contrasts exhibited enrichment of the tricarboxylic acid cycle TCA cycle (mmu00020) KEGG pathway and a number of GO MF terms linked to metabolism of cofactors and vitamins that are also linked to these metabolic pathways. These enriched categories are consistent with the higher metabolism of active muscles; elevated amino acid and energy metabolism are seen in muscles of

physically active mice, and presumably this elevated amino acid metabolism maintains the TCA cycle intermediates needed for fatty acid metabolism (Wone et al., 2011).

Muscle cell differentiation (GO:0042692) and muscle organ development (GO:0007517) were two BP terms enriched among the genes differentially expressed in the IW-IM and AM-AW contrasts. These categories are consistent with the known role of myostatin on cell differentiation and proliferation in triceps. Multiple studies have confirmed the direct impact of myostatin on these muscles. Specifically, myostatin-deficient mice have significantly larger tricep muscles than wild-type mice (Personius et al., 2010; Savage & McPherron, 2010; Hamrick et al., 2002). Myostatin depletion increased muscle mass by an average of 28%-44% in sedentary mice (Personius et al., 2010).

Enrichment of hypertrophic and dilated cardiomyopathy KEGG pathways (mmu05410 and mmu05414, respectively) was observed among genes differentially expressed in the IW-IM and AM-AW contrasts. Although hypertrophic cardiomyopathy is characterized by an hypertrophied heart muscle while tricep samples were used in this study, our results suggest that the expression of genes in similar biological processes are altered by myostatin genotype regardless of activity level. Our result is consistent with a previous report that a hypertrophic cardiomyopathy mutation is expressed in the messenger RNA of skeletal as well as cardiac muscle (Yu et al., 1993).

Finally, enrichment of the vasculature development BP (GO:0001944) among the genes differentially expressed in the AM-IM and IW-IM contrasts suggests that activity level within myostatin-reduced mice and myostatin genotype within inactive mice have comparable impact on the expression of genes in the vascular development pathway. While inactivity may have counteracted the effect of myostatin reduction in the former contrast, myostatin reduction may have counteracted the effect of inactivity in the latter contrast. Our results offer support at the gene expression level to claims that the processes that regulate blood vessel development can also enable the adult to adapt to changes in tissues that can be elicited by activity or pathologies (Udan et al., 2013).

Notable genes. Consideration of individual genes exhibiting significant activity-by-genotype interaction further augmented the understanding of the interplay between activity and myostatin genotype on the triceps brachii muscle transcriptome in C57/BL6

mice. PAK1 displayed consistent borderline significant differential expression across multiple contrasts, resulting in a significant overall interaction effect.. A study of target genes of myostatin loss-of-function in muscles of bovine fetuses identified PAK1 (Cassar-Malek et al., 2007). The complex interplay between PAK proteins in regulation of signaling cascades controlling cell motility, proliferation, and morphology and reorganization of the cytoskeleton (Schneeberger & Raabe, 2003) may lead to compensatory mechanisms resulting in wild-type mice exhibiting higher levels of PAK1 than myostatin-reduced mixed within the inactive group. Supporting this hypothesis, signaling of PAK1 has been linked to the G1 to S phase transition of the cell cycle via regulation of cyclin D1 machinery (Heng & Koh, 2010). This function is consistent with our findings of high levels of expression of MIR1945- G1 to S phase transition 1 (Gspt1) in our genotype contrast of myostatin-reduced vs. wild-type mice. Both genes are known to be associated with SMAD3 that in turn is associated with TGFB1. Similarly, myostatin negatively regulates the activation of satellite cells by controlling the G1 to S phase transition through down-regulation of Cdk2 and up-regulation of P21, the protein encoded by PAK1 (McCroskery et al., 2003). The expression of PAK1 was highest in active myostatin-reduced mice, followed by inactive wild-type mice and then inactive myostatin-reduced mice. The detection of this gene further confirms the importance of G1 to S phase transition 1 in muscular physiology and the role of myostatin in inhibition. PAK1 appears to be activated during the process of vascular remodeling (Hinokie et al., 2010) and this is in agreement with the identification of enrichment of vascular development pathway in the IM-AW and IW-IM contrasts.

CSRP3 and Myoz2 shared the same interaction pattern of highest expression in inactive myostatin-reduced relative to all other activity level genotype groups. The parallel expression profiles of these two genes detected in the present study is in agreement with previous reports. The expression of CSRP3 and Myoz2 is high in skeletal muscles (Arber et al., 1994; Takada et al., 2001), positively regulating myogenesis through promotion of myogenic differentiation (Kong et al., 1997). CSRP3 encodes the muscle LIM protein (MLP), a muscle specific protein expressed and located at the z-line (Knoll et al., 2002) which has been described as essential for myogenesis given its potential for induction of myogenic differentiation (Arber et al., 1994). Mice with a

deficiency of this gene exhibit dilated cardiomyopathy (Arber et al., 1997). Only a few proteins have been shown to interact directly with MLP: actin (Arber & Caroni, 1996), alpha-actinin (Harper et al., 2000), beta-spectrin (Flick & Konieczny, 2000), and N-RAP (Ehler et al., 2001); A definitive link between myostatin and MLP has not been established. Other studies have found a relationship between expression of MLP and contractility (Ecarnot-Laubriet et al., 2000; Schneider et al., 1999). In addition to playing structural and functional roles in skeletal muscle, MLP has been suggested to be a mediator of mechanical stress in cardiac tissue (Barash et al., 2005). Muscle growth resulting from myostatin inactivation presumably creates an imbalance between the metabolic requirements of tissue cells and the previous perfusion capabilities of blood vessels, and CSRP3-encoded MLP may work to mediate this stress and reduce likelihood of cardiomyopathy.

Clusters of expression profiles among genes exhibiting significant activity-by-genotype interaction were identified as well. Among genes sharing the first profile (under-expression in active wild-type and active myostatin-reduced relative to inactive myostatin-reduced and similar expression levels across all other activity-genotype groups), KEGG pathways for several inflammation-associated neurodegenerative conditions including Parkinson's disease, Alzheimer's disease, and Huntington's disease were enriched. Our results are in agreement with reports that myostatin causes sporadic inclusion body myositis (sIBM), a muscle-wasting disease that has pathogenesis similar to that of Alzheimer's and Parkinson's diseases (Starck & Sutherland-Smith, 2010). Also, activin A protects from neural degeneration in individuals with Huntington's disease, (Hughes et al., 1999) and the relationship between myostatin and activin has been well established. Myostatin signals muscle mass control through activin receptors (Lee, 2004), meanwhile activin type IIB receptor acts as a myostatin inhibitor, causing a dramatic muscle mass increase (Lee et al., 2005). In addition to the previous pathways, oxidation-reduction, oxidoreductase activity, and oxidase activity categories were also enriched among genes in the first profile. This enrichment is consistent with studies demonstrating that oxidative stress is often induced by physical activity due to the generation of reactive oxygen species (ROS) that occurs as skeletal muscles contract (Powers & Jackson, 2008). Also, myostatin acts as a pro-oxidant, inducing oxidative stress in skeletal muscle by

inducing ROS (Sriram et al., 2014). In turn, this induces anti-oxidant enzymes in skeletal muscle through TNF-α and NADPH oxidase in a feed-forward manner (Sriram et al., 2011). Additional GO terms associated with the electron transport chain, such as generation of precursor metabolites and energy, monovalent inorganic transmembrane transporter activity, and inorganic cation transmembrane transporter activity, enriched in the first profile highlight the expected link between muscle function and mitochondria-dependent reformation of ATP through nutrient oxidation.

Additionally, oxidative phosphorylation, electron transport chain, energy generation, and mitochondrial ATP synthesis GO terms were enriched among the profile characterized by genes under- or over-expressed in the AW-IM contrast and not differentially expressed in all other contrasts. These findings are consistent with the electron transport chain, or the flow of electrons resulting from NADH and FADH2 oxidation, that establishes an electrochemical gradient vital in powering ATP synthesis in oxidative phosphorylation, the final stage of aerobic cell respiration. Myostatin reduction, although not affecting phosphorylated compound concentrations and intracellular pH at rest, causes up to a 206% increase in ATP cost of contraction as well as limiting the shift toward oxidative metabolism during muscle activity (Giannesini et al., 2013). Muscle buildup caused by myostatin is sustained through a combination of reduced ATP synthesis and decreased protein degradation activity (Uyttenhove et al., 2003)

Antagonistic and synergistic expression patterns. Finally, using genes exhibiting significant (FDR-adjusted P-value < 2XE-12) genotype-by-activity interaction effects, antagonistic and synergistic expression patterns were identified. Synergistic effects occur when the expression of a gene under a combination of genotype and activity levels is more extreme than the average expression under each level separately. Antagonistic effects occur when the expression of a gene under a combination of genotype and activity levels is less extreme than the average expression under each level separately. An example of synergistic pattern would be when a gene that has high over-expression (e.g. 4 fold) in myostatin inactive relative to the average of all other groups meanwhile the expression in myostatin relative to wild type and the expression in inactive relative to active are less or non-significant (e.g. less than 2 fold). An example of antagonistic pattern would be a gene that is not or less differentially expressed (e.g. 1 fold) in

myostatin inactive mice relative to the average of all other groups meanwhile the expression in myostatin relative to wild type and the expression in inactive relative to active are more significant or extreme (e.g. more than 3 fold).

Examples of synergistic or antagonistic mode of action of genotype and activity factors on gene expression are listed in Table 2.1. Mettl21e; Cytochrome P450, Family 1, Subfamily A, Polypeptide 1 (Cyp1a1); and Myelin protein zero (Mpz) were identified as antagonistic genes, where gene expression increases in one contrast and concurrently decreases in another contrast. Cyp1a1 and Mpz shared the same interaction pattern, characterized by lower expression in active wild-type mice relative to the inactive wild-type mice, and by higher expression in active myostatin-reduced relative to inactive myostatin-reduced mice. A striking example of antagonistic interaction among these significant genes is Mettl21e, which follows the opposite interaction pattern. The expression of Mettl21e in active wild-type mice is higher than that in inactive wild-type mice, and the expression in active myostatin-reduced mice is lower than in inactive myostatin-reduced mice. Similarly, Sarcolipin (Sln), Actin Alpha 2 (Acta2), nuclear receptor corepressor 2 (Ncor2), guanine nucleotide binding protein beta polypeptide 2-like 1 (Gnb2l1), and Gremlin 2 of the Cysteine Knot Superfamily (Grem2) displayed antagonistic interactions (Table 2.1 and Table S2.1).

The identification of significant interactions enabled the detection of synergistic effects between genotype and activity. For genes Naca, Dusp23, and Dhcr24, the difference in expression between myostatin genotype groups was more extreme than between activity groups (Table S2.1). The changes in gene expression between genotypes appeared to be magnified by activity. This suggests that the observed changes in transcript abundances identified in the activity-level contrast may be due to the sole effect of physical activity, while changes identified in the genotype contrast may be due to the effect of both physical activity and myostatin depletion. The similar expression levels between genotypes among inactive mice and striking differential expression between genotypes among the active mice indicates strong synergistic interplay. Other genes also exhibiting synergistic interaction between genotype and activity, included Dimethylarginine Dimethylaminohydrolase 1 (Ddah1), FBJ Murine Osteosarcoma Viral Oncogene Homolog (Fos), WNK Lysine Deficient Protein Kinase 2 (Wnk2),

Transmembrane Protein 100 (Tmem100), Prostate Transmembrane Protein Androgen Induced 1 (Pmepa1), Receptor G Protein-Coupled Activity Modifying Protein 1 (Ramp1), LanC Lantibiotic Synthetase Component C-Like 1 (Lancl1), P21 Protein Cdc42/Rac-Activated Kinase 1 (Pak1), and Attractin-Like 1 (Atrnl1).

Understanding the main effect of myostatin on gene expression in triceps brachii muscles

Among the genes differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |1.3|) between myostatin-reduced and wild-type mice (Table 2.7 and Table S2.7), the functions of three genes are associated to muscle physiology. The detection of Actin, alpha, cardiac muscle 1 (Actc1) in this study is consistent with reports that cardiac α-actin can functionally substitute at least in part for skeletal muscle α-actin in skeletal muscle (Ravenscroft et al., 2008) Similarly, the detection of Sln differential expression between myostatin genotype groups is consistent with reports postulating that high Sln expression in human skeletal muscle is important to the physiology of the tissue (Fajardo et al., 2013). Interestingly, Sln has recently been shown to mediate muscle-based thermogenesis (Bal et al., 2012). Finally, differential expression of Grem2 has been detected in the quadricepts of a mouse model of non-dystrophic skeletal muscle congenital disease (Gineste et al., 2013). Also, Grem2 acts as an antagonist of bone morphogenetic proteins (BMPs) that influence the effectiveness of myostatin. Myostatin is synthesized as a precursor protein, which then becomes biologically active through BMP-driven proteolytic processing events (Lee, 2010; Wolfman et al., 2003). The profile observed in this study can be explained by myostatin gene expression depletion requiring lower BMP convertase, which is achieved through the inhibitory action of Grem2.

Among the rest of the genes differentially expressed between genotype groups, four provided remarkable insight into myostatin's effects on the gene networks of muscle development. The protein coded by the Piezo-type mechanosensitive ion channel component 1 (Piezo1) allows cells to react to physical stimuli. Mechanosensitive ion channels play a key role in the physiology of smooth muscle (Fajardo et al., 2013; Kirber et al., 1988; Davis et al., 1992). Consistent with the differential expression of XPD (also known as Excision repair cross-complementing rodent repair deficiency complementation

group 2 or Ercc2) between genotype groups, a mutation on XP genes has been associated with a reduction in skeletal muscle in mice (de Boer et al., 2002). Similarly to the present study, Microprocessor complex subunit DGCR8 (Dgcr8) or DiGeorge syndrome critical region gene 8 has been linked to myoblast differentiation (Kozakowska et al., 2012). Finally, our results revealed the expression of gene G1 to S Phase Transition 1 (Gspt1), responsible for the G1 to S phase transition of the cell cycle (Bartek & Lukas 2001). Given the differential expression pattern of the present study, we propose that that myostatin could inhibit the process of myoblasts moving from the G1 to S phase of the cell cycle through up-regulation of p21 and subsequent inhibition of Cdk2 activity. An extended list (FDR-adjusted P-value < 0.01**)** of differentially expressed genes between myostatin-reduced and wild-type mice in triceps brachii muscles can be found in Table S2.7.

Understanding the main effect of physical activity on gene expression in triceps brachii muscles

Many of the differentially expressed genes between active and inactive mice, unsurprisingly, are associated with the biological processes of contractile response of muscles to activity. A notable finding is that Ercc2 was differentially expressed between genotype groups and between activity groups as well, yet this gene did not exhibit a significant activity-by-genotype interaction effect. A similar molecular mechanism is speculated for both comparisons. Among other genes differentially expressed between activity groups, Tnnt1 (Samson et al., 1992) plays an essential role in skeletal muscle contraction by regulating calcium sensitivity (Nadal-Ginard & Mahdavi, 1989). The myoglobin protein, encoded by Mb, is present only in myocytes and oxidative skeletal muscle fibers. This gene is essential for oxygen storage in muscle (Ordway & Garry, 2004), and facilitates oxygen diffusion by desaturating rapidly as muscle activity increases (Wittenberg, 1970). Cyp26b1 encodes protein Cytochrome P450 26B1, known to be present in adult mice skeletal muscle (Choudhary et al., 2005). Cyp26b1 signals aortic smooth muscle cells through regulation of the metabolism of all-trans-retinoic acid, (Ocaya et al., 2010) which is crucial for regulation of gene expression, cell growth and

differentiation (Ateia Elmabsout et al., 2012). Finally, the protein encoded by Tropomyosin alpha-3 (TPM3) is also essential for regulation of skeletal muscle contraction (Morris et al., 1991).

Novel associations between differentially expressed genes and activity level were also identified in this study. Many of these genes have indirect links to muscle function and activity, but the actual mechanism uncovered is unique and unexpected. The differentially expressed gene 3-hydroxybutyrate dehydrogenase (Bdh1) encodes an enzyme involved in the interconversion of acetoacetate and (R)-3-hydroxybutyrate, essential for fatty acid catabolism. Also, Bdh1 mRNA is found in all forms of muscle (Marks et al., 1992). The role in catabolism could be associated with the need for energy that characterizes the skeletal muscle under activity. Regarding the enrichment of the ATP metabolic process, physical activity causes an increase in ATP cost of contraction (Giannesini et al., 2013). At the same time, active mice have significant reduction in accumulation of body fat as compared to wild-type IQ motif Sec7 domain 2 (Iqsec2 or Brag2) has been associated with myoblast cell-cell fusion (Pajcini et al., 2008). In addition, estrogen-related receptor beta (ERRbetta) is a nuclear receptor protein encoded by the Esrrβ gene that was differentially expressed among activity groups. This result is consistent with work demonstrating that ERRbeta/gamma agonist modulates GRalpha expression, and glucocorticoid responsive gene expression in skeletal muscle cells (Wang et al., 2010). Finally, AK4 was differential expressed between activity groups and this gene is responsible for encoding adenylate kinase 4, an energy-mediating enzyme. This finding is in agreement with reports that AK4 is highly expressed in human skeletal muscle (Noma et al., 2001).

The enrichment of GO biological process terms related to vasculature development (angiogenesis, blood vessel development, vasculature development, blood vessel morphogenesis, and tube development) among the genes differentially expressed in the AM-IM and IW-IM contrasts suggests that the combination of activity and myostatin-reduced genotype has comparable impact to the combination of inactivity and wild-type typical myostatin genotype on the expression of genes in the vascular development pathway. A link between vascular development and muscle development is expected based on the logical physiological association of the two organ systems.

Vasculature is modified in order to meet the metabolic requirements of tissue cells in response to changes in metabolic rate; oxygen is a major control element of this adaptation, as hypoxia initiates various signals which in turn lead to an increase in vessel growth (Adair et al., 1990). Given this information, vascular growth and activation of vascular development pathways would be expected upon myostatin inactivation, as the resulting muscle growth presumably creates an imbalance between the metabolic requirements of tissue cells and the previous perfusion capabilities of blood vessels. In addition, genes activated during vascular processes, such as PAK1, were found through our analyses to exhibit significant activity-by-genotype interaction. The association between myostatin level and PAK1 was confirmed in a previous study of genes targeted by myostatin loss-of-function in bovine muscles (Cassar-Malek et al., 2007).

CONCLUSION

The study of the impact of physical activity and myostatin level on gene expression in the triceps brachii muscles of C57/BL6 mice uncovered novel and confirmed known associations at the gene and gene network levels. Novel and significant interaction effects were observed for some genes (e.g. Naca, Grem2) including synergistic effects (e.g. Naca, Dhcr24) and antagonistic effects (e.g. Mettl21e, Cyp1a1, Mpz). Functional analysis of genes presenting significant interaction effects uncovered novel (e.g. angiogenesis) and expected (e.g. oxidative phosphorilation, electron transport chain) enriched pathways and biological processes.

Among the genes exhibiting significant main genotype effect, known (e.g. Sln, Grem2) and novel (e.g. Piezo1, Ercc2, Gspt1) associations were detected. Functional analysis of genes presenting significant genotype effect uncovered novel (e.g. dilated and hypertrophic cardiomyopathy) and expected (e.g. muscle cell differentiation, muscle organ development) enriched pathways and biological processes. Likewise, among the genes exhibiting significant main activity effect, known (e.g. MB, Tpm3) and novel (e.g. Bdh1, Esrrβ) associations were detected. Functional analysis of genes presenting significant activity effect uncovered novel (e.g. Alzheimer's, Parkinson's and Huntington's disease) and expected (e.g. oxidative phosphorylation, cardiac muscle

contraction) enriched pathways and biological processes. While several genes and functional categories enriched among the differentially expressed genes uncovered in this study were consistent with previous reports, the identity and profile of the genes exhibiting the most extreme interaction and main genotype and activity effects opened new avenues of inquiry on the role of specific genes in skeletal muscle development and the effects of myostatin and physical activity on muscle function. The present study centered on the comparison of four genotype-activity groups based on transcriptome information from a specific skeletal muscle type, mouse strain, gender, and age. Consideration of additional muscle types, genotypes, activities, ages, and genders would help identify additional synergistic and antagonistic relationships between these factors.

The findings from the present study could have medical implications on preventive practices and therapies associated with muscle atrophy in humans and companion animal species and genome-enabled selection practices applied to food-production animal species. The study of changes in gene expression in response to myostatin gene expression level in skeletal muscle tissue involved genes that code for a number of proteins that are feedback regulated by the myostatin molecule. The functions of the genes exhibiting differential expression between genotype groups are primarily regulatory. This functional category includes microRNA and Piezo proteins that make the list of the top 10 differentially expressed genes, side-by-side with Grem2 proteins that modulate the metalloprotein BMP-mediated cleavage of the myostatin propeptide. The role of genes regulated by microRNAs was unanticipated, especially because these genes seem to impact central functions such as the G1 to S phase transition of the cell cycle of myoblasts. The role of genes coding for Piezo proteins that make mechanosensitive ion channels, which in turn regulate cationic currents in the cells, was also remarkable and unanticipated, especially because of the consistent profile of the genes in this family. The study of changes in gene expression patterns in response to activity level revealed enrichment of genes that code structural proteins important for muscle function, including troponin, tropomyosin and myoglobin proteins. Activity was also associated with differential expression of genes important for fatty acid metabolism, some linked to type II diabetes and obesity and others to DNA-repair capacity, stem cell renewal, and various forms of cancer.

Our results provide evidence supporting the role of myostatin as a master regulator and the hypothesis that physical activity affect the expression of genes associated with homeostatic balance between storage of fat and muscle growth. Down-regulation of myostatin expression enables muscle growth at full expense of storage of fat, a condition that is hardwired at the regulatory level (e.g. through antagonists of metalloenzymes responsible for the myostatin activation). During activity, the changes in gene expression associated with balance between storage of fat and growth appears more instantaneous and subtle. This balance involves the regulation of metabolic pathways of fatty acid synthesis and does not impinge on oxidative phosphorylation pathways. The master regulatory functions of myostatin identified in this study should now be explored at the biochemical level to identify details of the regulatory networks, especially because of their potential to assist in the development of muscular disorders.

TABLES AND FIGURES

Table 2.1 Differentially expressed genes (FDR-adjusted P-value < 2XE-12) across activity-genotype contrasts

Gene Name*	Log2Fold[1]						FDR-adjusted P-Value
	AM-AW[3]	AM-IM[4]	AW-IM[2]	IW-IM[7]	IW-AW[5]	IW-AM[6]	
Mettl21e	*2.72*	*0.57*	*3.29*	*4.48*	*-1.19*	*-3.91*	*1.55E-13*
Dusp18	-0.78	-0.73	-1.51	-0.64	-0.87	-0.09	1.55E-13
Per1	1.22	0.40	1.62	0.82	0.79	-0.42	1.55E-13
Atp1b2	0.62	0.34	0.96	0.31	0.65	0.023	1.55E-13
Tnnc1	-0.89	-4.14	-5.03	-3.67	-1.36	-0.47	1.55E-13
Zmynd17	-1.25	0.47	-0.78	-1.63	0.85	2.09	1.55E-13
Myh7	-0.72	-4.87	-5.58	-4.37	-1.20	-0.49	1.55E-13
Tpm3	-0.75	-2.22	-2.96	-1.76	-1.20	-0.46	1.55E-13
Ddah1	*-1.64*	*0.08*	*-1.57*	*-2.15*	*0.59*	*2.23*	*1.55E-13*
Myl2	-1.15	-3.74	-4.89	-3.49	-1.39	-0.25	1.55E-13
Atp2a2	-0.89	-2.15	-3.05	-2.31	-0.74	0.16	1.55E-13
Pak1	*0.56*	*-0.09*	*0.46*	*1.07*	*-0.60*	*-1.16*	*1.55E-13*
Tnnt1	-0.92	-4.07	-4.99	-3.66	-1.33	-0.41	1.55E-13
Csrp3	-0.91	-1.92	-2.83	-2.01	-0.82	0.09	1.55E-13

Table 2.1 (cont.)

Fxyd6	-0.76	-2.05	-2.81	-1.96	-0.85	-0.08	1.55E-13
Myoz2	-0.59	-2.43	-3.03	-2.12	-0.91	-0.31	6.84E-13
Myl3	-0.62	-3.15	-3.77	-2.76	-1.01	-0.39	6.84E-13
Naca	0.87	0.03	0.91	0.43	0.48	-0.39	1.89E-12
Ak3	0.55	-1.53	-0.99	0.41	-1.40	-1.95	1.89E-12
Cyp1a1	2.0428	-0.8460	1.19676	0.26727	0.929775	-1.1131	3.16E-12
Grem2	*1.44986*	*0.213147*	*1.66277*	*2.52511*	*-0.86228*	*-2.31217*	*3.16E-12*
Fos	*-0.78073*	*0.039794*	*-0.74082*	*-1.52664*	*0.786004*	*1.56659*	*8.48E-12*
Gnb2l1	0.739797	-0.36408	0.375364	-0.32355	0.699348	-0.04009	8.56E-12
Wnk2	1.10552	0.056718	1.16275	0.760786	0.402344	-0.70346	2.45E-11
Tmem100	*-0.76578*	*0.050259*	*-0.71569*	*-1.11695*	*0.401146*	*1.16691*	*3.02E-11*
Ncor2	0.556591	-0.14625	0.41082	-0.02854	0.439626	-0.11714	6.96E-11
Acta2	0.972604	-0.62395	0.348594	-0.39031	0.738874	-0.23374	2.01E-10
Pmepa1	*-0.67603*	*-0.0258*	*-0.70168*	*-1.02573*	*0.324245*	*1.00005*	*2.01E-10*
Sln	*1.5134*	*0.156663*	*1.67004*	*2.36704*	*-0.69664*	*-2.21015*	*2.90E-10*

Table 2.1 (cont.)

Dhcr24	0.771441	-0.05624	0.715401	-0.12054	0.83638	0.064728	3.20E-10
Dusp23	0.468437	0.024808	0.493519	0.026371	0.467286	-0.00141	1.01E-09
Ramp1	-0.66605	-0.05061	-0.71666	-0.92948	0.213069	0.87899	1.07E-09
Atrnl1	-0.67904	-0.04955	-0.72883	-0.49233	-0.2365	0.442585	1.32E-09
Lancl1	*-1.22412*	*0.08892*	*-1.13543*	*-1.33056*	*0.195232*	*1.41926*	*1.43E-09*

* Genes exhibiting significant synergistic and antagonist activity-by-genotype interaction effects are displayed in italics. Ddha1, Lancl1, Fos, Tmem1, and Pmepa1 follow a synergistic pattern; Sln, Grem2, Mettl2, and Pak1 follow an antagonistic pattern.

1. When considering two values, A and B, Log2Fold Change = Log2 (B/A). For example, Log2Fold of the contrast AW-IM = log2(IM/AW).
2. AW-IM refers to the active wild-type vs. inactive myostatin-reduced contrast group
3. AM-AW refers to the active myostatin-reduced vs. active wild-type contrast group
4. AM-IM refers to the active myostatin-reduced vs. inactive myostatin-reduced contrast group
5. IW-AW refers to the inactive wild-type vs. active wild-type contrast group
6. IW-AM refers to the inactive wild-type vs. active myostatin-reduced contrast group
7. IW-IM refers to the jnactive wild-type vs. inactive myostatin-reduced contrast group

*Expanded gene names, listed in alphabetical order: Acta2= actin, alpha 2, smooth muscle, aorta; Ak3= adenylate kinase 3; Atp1b2= ATPase, Na+/K+ transporting, beta 2 polypeptide; Atp2a2= ATPase, Ca++ transporting, cardiac muscle, slow twitch 2; Atrnl1= Attractin-Like 1; Csrp3= cysteine and glycine-rich protein 3; Cyp1a1= Cytochrome P450, Family 1, Subfamily A, Polypeptide 1; Ddah1= dimethylarginine dimethylaminohydrolase 1; Dhcr24= 24-Dehydrocholesterol Reductase; Dusp18= dual specificity phosphatase 18; Dusp23= Dual Specificity Phosphatase 23; Fos= FBJ Murine Osteosarcoma Viral Oncogene Homolog; Fxyd6= FXYD domain-containing ion transport regulator 6; Gnb2l1= Guanine Nucleotide Binding Protein (G Protein), Beta Polypeptide 2-Like; Grem2= Gremlin 2, DAN Family BMP Antagonist; Lancl1= LanC Lantibiotic Synthetase Component C-Like 1; Mettl21e= methyltransferase like 21E; Myh7= myosin, heavy chain 7, cardiac muscle, beta; Myl2= myosin, light polypeptide 2, regulatory, cardiac, slow; Myl3= myosin, light polypeptide 3; Myoz2= myozenin 2; Naca= nascent polypeptide-

Table 2.1 (cont.)

associated complex alpha polypeptide; Ncor2= Nuclear Receptor Corepressor 2; Pak1= p21 protein (Cdc42/Rac)-activated kinase 1

Per1= period circadian clock 1; Pmepa1= Prostate Transmembrane Protein, Androgen Induced 1; Ramp1= Receptor (G Protein-Coupled) Activity Modifying Protein 1; Sln= Sarcolipin; Tmem100= Transmembrane Protein 100; Tnnc1= troponin C, cardiac/slow skeletal; Tnnt1= troponin T1, skeletal, slow; Tpm3= tropomyosin 3, gamma; Wnk2= WNK Lysine Deficient Protein Kinase 2; Zmynd17= zinc finger, MYND-type containing 17

Table 2.2 Shared differentially expressed genes (FDR-adjusted P-value < .01) between three orthogonal contrasts including the Inactive-Wild Type (IW) baseline group

Gene Name*	Log2(Fold)[1]		
	IW-AW[2]	IW-AM[3]	IW-IM[4]
M6prbp1	-0.55	-0.92	0.47
Pak1	-0.60	-1.16	1.07
Casq2	-0.87	-1.25	-0.75
Fos	0.79	1.57	-1.53
Dhrs4	-0.63	-0.64	-0.60
Gm5514	-1.44	-1.21	-1.52
Ddah1	0.59	2.23	-2.15
Fabp3	-0.75	-0.63	-0.93
Got1	-0.62	-0.55	-0.44
2310076L09Rik	-0.87	-0.72	-1.04
Mafb	-0.78	-1.07	0.58
EG225594	-1.64	-1.55	2.22
4832428D23Rik	-1.20	-3.91	4.48
BDH1	-1.83	-1.47	-2.24
ZMYND11	0.85	2.10	-1.63
Gck	1.17	1.48	-0.70
Esrrb	-1.26	-1.28	-1.28
Acaa2	-0.59	-0.53	-0.69
Ankrd2	-1.18	-0.92	-1.59
Actn2	-0.62	-0.78	-0.87
Egr1	1.56	1.65	-1.48
Myom3	-1.13	-1.08	-1.93
9830123M21Rik	0.69	2.52	-0.49

Table 2.2 (cont.)

Rn45s	-0.80	-1.94	1.16
NNT	-0.79	-0.72	-0.70
Dgat2	-0.83	-0.70	-0.97
H19	-0.68	-0.89	0.43
Tbc1d1	0.72	0.98	-0.91
IL15	-1.05	-1.39	1.06
Myh2	-1.08	-0.85	-1.78
COL22A1	0.63	0.74	-0.48
ORF63	0.85	0.93	-0.75
IDH2	-0.85	-0.67	-0.92
M6prbp1	-0.55	-0.92	0.47

1. When considering two values, A and B, Log2Fold Change = Log2 (B/A). For example, Log2Fold of the contrast AW-IM = log2(IM/AW).

5. IW-AW refers to the inactive wild-type vs. active wild-type contrast group

6. IW-AM refers to the inactive wild-type vs. active myostatin-reduced contrast group

7. IW-IM refers to the jnactive wild-type vs. inactive myostatin-reduced contrast group

*Expanded gene names, listed in alphabetical order: 2310076L09Rik= RIKEN cDNA 2310076L09 gene; 4832428D23Rik=RIKEN cDNA 4832428D23 gene; 9830123M21Rik= RIKEN cDNA 9830123M21 gene; Acaa2= acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme A thiolase); Actn2=actinin alpha 2; Ankrd2=ankyrin repeat domain 2 (stretch responsive muscle); BDH1= 3-hydroxybutyrate dehydrogenase, type 1; Casq2=calsequestrin 2; COL22A1= collagen, type XXII, alpha 1; Ddah1=dimethylarginine dimethylaminohydrolase 1; Dgat2=diacylglycerol O-acyltransferase 2; Dhrs4= dehydrogenase/reductase (SDR family) member 4; EG225594=predicted gene 4841; Egr1=early growth response 1; Esrrb=estrogen related receptor, beta; Fabp3=fatty acid binding protein 3, muscle and heart; similar to mammary-derived growth inhibitor; Fos= FBJ osteosarcoma oncogene; Gck= glucokinase;Gm5514= lactate dehydrogenase B; predicted gene 5514; Got1= similar to Aspartate aminotransferase, cytoplasmic (Transaminase A) (Glutamate oxaloacetate transaminase 1); glutamate oxaloacetate transaminase 1, soluble; H19= H19 fetal liver mRNA; IDH2= Isocitrate Dehydrogenase 2 (NADP+), Mitochondrial; IL15=interleukin 15; M6prbp1= mannose-6-phosphate receptor binding protein 1; Mafb=v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B (avian) ;Myh2= myosin, heavy polypeptide 2, skeletal muscle, adult, myosin, heavy polypeptide 1, skeletal muscle, adult; Myom3=myomesin family, member 3; Nnt= nicotinamide nucleotide transhydrogenase; ORF63= open reading frame 63; Pak1= p21 protein (Cdc42/Rac)-activated kinase 1; Rn45s= RNA, 45S Pre-Ribosomal 5

Table 2.3 Most common profiles of over- (denoted with +1), under- (denoted with -1) and not (denoted with 0) differentially expressed (P-value < 0.0001) genes across the 6 pairwise contrasts

Profile[1]						**Number of Genes**[2]
AM-AW	AM-IM	AW-IM	IW-IM	IW-AW	IW-AM	
0	0	0	0	0	0	1509
0	0	+1	0	0	0	183
0	0	-1	0	0	0	146
0	-1	-1	0	0	0	142
0	-1	0	0	0	0	86
0	+1	0	0	0	0	83
0	0	+1	+1	0	0	71
0	+1	+1	0	0	0	68
0	-1	-1	-1	0	0	67
0	+1	+1	+1	0	0	54

1. The six contrast groups are ordered as follows- active wild-type vs. inactive myostatin-reduced, active myostatin-reduced vs. active wild-type, active myostatin-reduced vs. jnactive myostatin-reduced, inactive wild-type vs. active wild-type, inactive wild-type vs. active myostatin-reduced, and jnactive wild-type vs. inactive myostatin-reduced. Each number of the 6 in the Profile "code" refers to each contrast in order (the first number in the Profile "code" denotes significance level for the first group, active wild-type vs. inactive myostatin-reduced, the second number denotes significance level for the second group, etc.)

2. Unlisted profiles include < 50 genes

Table 2.4 Enriched (enrichment score > 3.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes exhibiting the profile characterized by under-expressed in active wild-type vs inactive myostatin-reduced and in active myostatin-reduced vs inactive myostatin-reduced and not differentially expressed in all other contrasts.

Category	**Term**	**Number of Genes**	**P-Value**	**FDR-adjusted P-value**
Score = 16.28				
KEGG PATHWAY	mmu00190:Oxidative phosphorylation	31	6.69E-33	6.84E-30
KEGG PATHWAY	mmu05012:Parkinson's disease	30	6.01E-31	6.14E-28
GOTERM BP FAT	GO:0022900~electron transport chain	24	1.58E-27	2.39E-24
KEGG PATHWAY	mmu05010:Alzheimer's disease	30	9.82E-27	1.00E-23
KEGG PATHWAY	mmu05016:Huntington's disease	30	1.16E-26	1.19E-23
GOTERM BP FAT	GO:0006091~generation of precursor metabolites and energy	29	2.84E-25	4.30E-22
GOTERM BP FAT	GO:0055114~oxidation reduction	32	3.95E-17	5.99E-14

Table 2.4 (cont.)

GOTERM MF FAT	GO:0015078~hydrogen ion transmembrane transporter activity	12	1.27E-11	1.65E-08
GOTERM MF FAT	GO:0015077~monovalent inorganic cation transmembrane transporter activity	12	2.47E-11	3.20E-08
GOTERM MF FAT	GO:0022890~inorganic cation transmembrane transporter activity	12	1.55E-09	2.01E-06
KEGG PATHWAY	mmu04260:Cardiac muscle contraction	10	2.24E-07	2.29E-04
GOTERM MF FAT	GO:0015002~heme-copper terminal oxidase activity	6	5.10E-07	6.61E-04
GOTERM MF FAT	GO:0016675~oxidoreductase activity, acting on heme group of donors	6	5.10E-07	6.61E-04
GOTERM MF FAT	GO:0016676~oxidoreductase activity, acting on heme group of donors, oxygen as acceptor	6	5.10E-07	6.61E-04
GOTERM MF FAT	GO:0004129~cytochrome-c oxidase activity	6	5.10E-07	6.61E-04

Table 2.5 Enriched (enrichment score > 3.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes under- or over-expressed in the active wild-type vs inactive myostatin-reduced contrast and not differentially expressed in all other contrasts.

Category	**Term**	**Number of Genes**	**P-Value**	**FDR-adjusted P-value**
Score = 5.07				
GOTERM BP FAT	GO:0006091~generation of precursor metabolites and energy	24	1.74E-10	2.86E-07
GOTERM BP FAT	GO:0022900~electron transport chain	15	8.68E-09	1.43E-05
GOTERM BP FAT	GO:0055114~oxidation reduction	32	6.91E-07	1.1E-03
GOTERM BP FAT	GO:0045333~cellular respiration	9	7.84E-06	0.01
GOTERM BP FAT	GO:0015980~energy derivation by oxidation of organic compounds	11	7.88E-06	0.01
GOTERM BP FAT	GO:0022904~respiratory electron transport chain	6	1.10E-04	0.18
GOTERM BP FAT	GO:0006119~oxidative phosphorylation	7	4.14E-04	0.68

Table 2.5 (cont.)

Score = 4.18				
GOTERM BP FAT	GO:0006091~generation of precursor metabolites and energy	24	1.74E-10	2.86E-07
KEGG PATHWAY	mmu00190:Oxidative phosphorylation	16	7.12E-09	8.25E-06
GOTERM BP FAT	GO:0022900~electron transport chain	15	8.68E-09	1.43E-05
KEGG PATHWAY	mmu05016:Huntington's disease	18	1.82E-08	2.11E-05
KEGG PATHWAY	mmu05012:Parkinson's disease	15	7.73E-08	8.96E-05
KEGG PATHWAY	mmu05010:Alzheimer's disease	16	6.59E-07	7.63E-04
GOTERM BP FAT	GO:0006119~oxidative phosphorylation	7	4.14E-04	0.68

Table 2.6 Enriched (enrichment score > 2.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes under- or over-expressed in the active myostatin-reduced vs inactive myostatin-reduced contrast and not differentially expressed in all other contrasts.

Category	**Term**	**Number of Genes**	**P-Value**	**FDR-adjusted P-value**
Score = 2.43				
GOTERM BP FAT	GO:0001525~angiogenesis	8	3.00E-04	0.48

Table 2.7 Genes differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |1.3|) between myostatin-reduced and wild-type mice in the skeletal muscle.

Gene*	**Log2 (Myostatin-reduced/Wild-type)**	**FDR-adjusted P-value**
ERCC2	4.19	2.5E-03
DGCR8	4.09	2.5E-03
METTL21E	3.48	2.5E-03
GSPT1	2.38	2.5E-03
ACTC1	2.29	2.5E-03
GREM2	1.91	2.5E-03
SLN	1.89	2.5E-03
CDH4	1.88	2.5E-03
F830016B08RIK	1.87	2.5E-03
KATNAL2	1.74	2.5E-03
IL12A	1.47	2.5E-03
MYBPH	1.34	2.5E-03
VASH2	1.30	2.5E-03

*Expanded gene names, listed in alphabetical order: ACTC1=actin, alpha, cardiac muscle 1; CDH4=cadherin 4, type 1, R-cadherin; DGCR8=DGCR8 microprocessor complex subunit; ERCC2=Excision Repair Cross-Complementing Rodent Repair Deficiency, Complementation Group 2; F830016B08RIK=RIKEN cDNA F830016B08 gene; GREM2=gremlin 2, DAN family BMP antagonist; GSPT1=G1 to S phase transition 1; IL12A=interleukin 12A (natural killer cell stimulatory factor 1, cytotoxic lymphocyte maturation factor 1, p35); KATNAL2=katanin p60 subunit A-like 2; METTL21E=methyltransferase like 21E; MYBPH=myosin binding protein H; SLN=sarcolipin; VASH2=vasohibin 2

Table 2.8 Enriched (enrichment score > 3.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes differentially expressed between myostatin-reduced and wild-control mice (FDR-adjusted P-value < 0.05)

Category	**Term**	**Number of Genes**	**P-Value**	**FDR-adjusted P-value[1]**
Score = 3.59				
KEGG PATHWAY	mmu05410:Hypertrophic cardiomyopathy (HCM)	10	7.69E-05	0.09
KEGG PATHWAY	mmu04260:Cardiac muscle contraction	9	2.65E-04	0.31
KEGG PATHWAY	mmu05414:Dilated cardiomyopathy	9	8.16E-04	0.96
Score = 3.43				
GO BP FAT	GO:0007167~enzyme linked receptor protein signaling pathway	20	7.05E-06	0.01
GO BP	GO:0007179~transforming growth factor beta receptor signaling pathway	7	5.38E-04	0.89

[1] False Discovery Rate adjusted P-value. Only terms with FDR-adjusted P-value < 0.99 or with more than 5 genes are listed

Table 2.9 Genes differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |1.3|) between active and inactive mice in the skeletal muscle

Gene*	Log2(Active/Inactive)	FDR-adjusted P-value
ERCC2	3.95	1.54E-03
BDH1	2.38	1.54E-03
GM1078	2.28	1.54E-03
BC048679	2.28	1.54E-03
LRRC52	1.92	1.54E-03
LDHB	1.88	1.54E-03
TNNC1	1.86	1.54E-03
EGLN3	1.82	1.54E-03
MYL2	1.82	1.54E-03
TNNT1	1.78	1.54E-03
MYH7	1.78	1.54E-03
MYOM3	1.76	1.54E-03
ESRRB	1.74	1.54E-03
SLC26A10	1.65	1.54E-03
FHL2	1.65	1.54E-03
ANKRD2	1.63	1.54E-03
MYH2	1.61	1.54E-03
TM6SF1	1.59	1.54E-03
IQSEC2	1.57	1.54E-03
VAV2	1.53	1.54E-03
TPM3	1.51	1.54E-03

*Expanded gene names, listed in alphabetical order: ANKRD2=ankyrin repeat domain 2 (stretch responsive muscle); BC048679=cDNA sequence BC048679; BDH1=3-hydroxybutyrate dehydrogenase, type 1; EGLN3=egl-9 family hypoxia-inducible factor 3; ERCC2=excision repair cross-complementing rodent repair deficiency, complementation group 2; ESRRB=estrogen-related receptor beta; FHL2=four and a half LIM domains 2; GM1078=SH3 domain binding kinase family, member 3; IQSEC2=IQ motif and Sec7 domain 2; LDHB=lactate dehydrogenase B; LRRC52=leucine rich repeat containing 52;

Table 2.9 (cont.)

MYH2=myosin, heavy chain 2, skeletal muscle, adult; MYH7=myosin, heavy chain 7, cardiac muscle, beta; MYL2=myosin, light chain 2, regulatory, cardiac, slow; MYOM3=myomesin 3; SLC26A10=solute carrier family 26, member 10; TM6SF1=transmembrane 6 superfamily member 1; TNNC1=troponin C type 1 (slow); TNNT1=troponin T type 1 (skeletal, slow); TPM3=tropomyosin 3; VAV2=vav 2 guanine nucleotide exchange factor

Table 2.10 Enriched (enrichment score > 8.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes differentially expressed between active and inactive mice (FDR-adjusted P-value < 0.05)

Category	Term	Number of Genes	P-Value	FDR-adjusted P-value
Score = 19.79				
KEGG PATHWAY	mmu00190:Oxidative phosphorylation	51	5.29E-29	6.39E-26
KEGG PATHWAY	mmu05012:Parkinson's disease	50	2.05E-27	2.47E-24
GOTERM BP FAT	GO:0006091~generation of precursor metabolites and energy	65	2.35E-27	4.12E-24
KEGG PATHWAY	mmu05010:Alzheimer's disease	55	7.43E-25	8.98E-22
KEGG PATHWAY	mmu05016:Huntington's disease	53	5.65E-23	6.82E-20
GOTERM BP FAT	GO:0022900~electron transport chain	39	4.30E-22	7.54E-19
GOTERM MF FAT	GO:0015078~hydrogen ion transmembrane transporter activity	23	4.07E-11	6.28E-08
GOTERM MF FAT	GO:0015077~monovalent inorganic cation transmembrane transporter activity	23	1.44E-10	2.22E-07
GOTERM MF FAT	GO:0022890~inorganic cation transmembrane transporter activity	26	2.46E-09	3.79E-06

Table 2.10 (cont.)

Score = 8.42				
KEGG PATHWAY	mmu04260:Cardiac muscle contraction	28	1.13E-14	1.35E-11
GOTERM MF FAT	GO:0015078~hydrogen ion transmembrane transporter activity	23	4.07E-11	6.28E-08
GOTERM MF FAT	GO:0015077~monovalent inorganic cation transmembrane transporter activity	23	1.44E-10	2.22E-07
GOTERM MF FAT	GO:0022890~inorganic cation transmembrane transporter activity	26	2.46E-09	3.79E-06
GOTERM MF FAT	GO:0016675~oxidoreductase activity, acting on heme group of donors	10	7.32E-07	1.13E-03
GOTERM MF FAT	GO:0016676~oxidoreductase activity, acting on heme group of donors, oxygen as acceptor	10	7.32E-07	1.13E-03
GOTERM MF FAT	GO:0015002~heme-copper terminal oxidase activity	10	7.32E-07	1.13E-03
GOTERM MF FAT	GO:0004129~cytochrome-c oxidase activity	10	7.32E-07	1.13E-03

Figure 2.1 Basic RNA-sequencing protocol from extraction to analysis. The workflow that is boxed is repeated for each biological endpoint being analyzed.

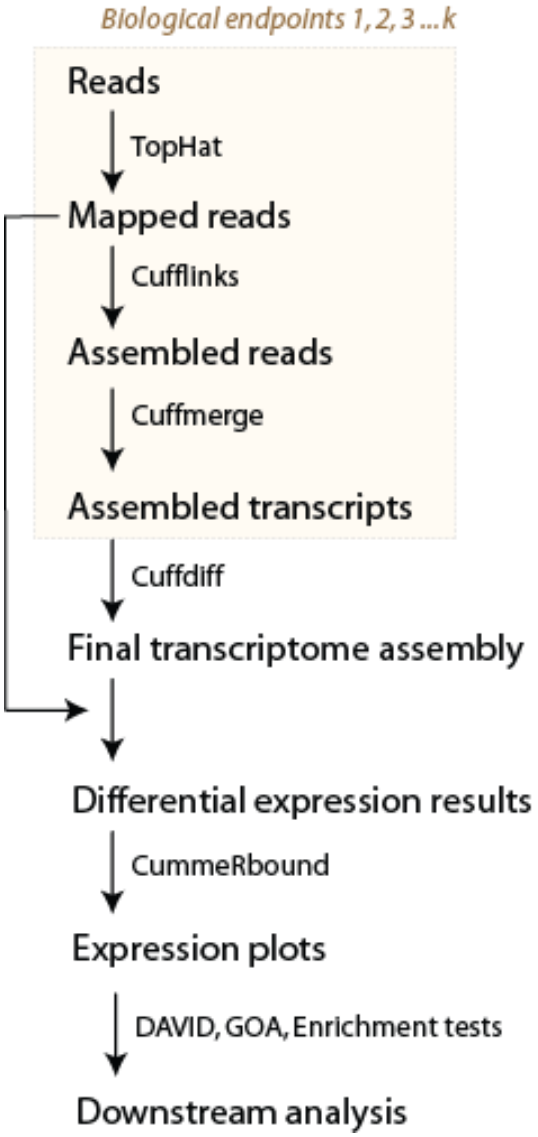
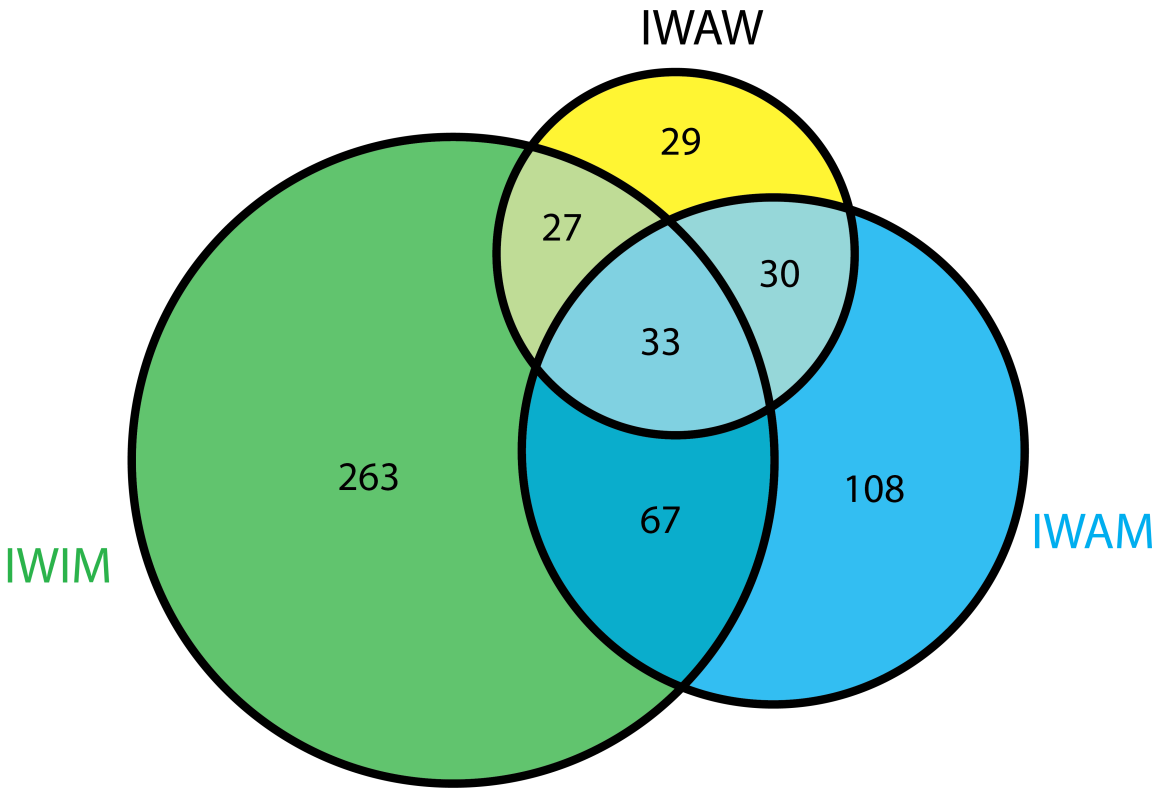


Figure 2.2 Number of differentially expressed genes that overlap between inactive wild-type (IW) and remainder activity-genotype groups: inactive myostatin-reduced (IM), active wild-type (AW), and active myostatin-reduced (AM).



CHAPTER 3. CEREBELLUM TRANSCRIPTOME OF MICE BRED FOR HIGH ACTIVITY OFFERS INSIGHTS INTO REWARD-DEPENDENT BEHAVIOR

Kelsey Caetano-Anollés[1], Justin S. Rhodes[2,3], Theodore Garland Jr[4], Sandra L. Rodriguez-Zas[1, 5, 6]*

Abstract

Addiction to activity and to psychoactive substances share psychologically and physiological characteristics. A transcriptome study compared a mouse line selected for high voluntary activity to a control line in environments restricting or enabling a rewarding activity. Results offered insights into genetic factors behind activity dependence and provided a model for understanding addiction and other reward-dependent behaviors.

Most genes differentially expressed between activity genotypes were only moderately differentially expressed between activity environments, suggesting that environmental effects were not confounded with activity genotype effects. Adora2a had a significant genotype-by-environment interaction effect evidenced by over-expression in the activity genotype relative to control in high activity environment and under-expression in the low activity environment. Our findings of differentially expressed genes related to dopaminergic transmission between activity genotypes support the association between these genes and activity. A central theme from the functional analysis of activity genotype-environment was neuron morphogenesis. Gene network analysis identified connected genes that exhibited similar (e.g. Lhx9) and opposite (e.g. Nrgn) expression patterns between activity genotypes across reward availability environment.

Our findings suggest that some transcriptomic changes in mice selected for high voluntary activity are shared with other reward-dependent behaviors. Results from this study support that high voluntary activity selection lines in mice are a helpful model to understand molecular mechanisms behind addition. Also, identification of genes and biological processes associated with both high voluntary activity and the pleasurable neurological response to physical activity may allow for the development of drugs which make it more pleasurable for people to exercise or less pleasurable to be sedentary as a

treatment approach for overweight/obesity as well as a way to improve health overall in the general population.

Introduction

The brain's dopaminergic mesocorticolimbic system has a primary role in reward-motivated response including associative learning or reinforcement (R. R. Brown et al., 1991; Ivlieva, 2011; Koob, 1992). Neurons in this region synapse and release dopamine in response to the introduction of addictive stimuli (Hyman, 2005; Kalivas & Volkow, 2005; McKim, 2003; Vivier, Tomasello, Baratin, Walzer, & Ugolini, 2008). Psychoactive substances, including opiates, cocaine, amphetamine, alcohol, and nicotine impact the neurotransmission of dopamine just as other more natural rewards including specific foods or liquids, and exercise (Bowman, Aigner, & Richmond, 1996; Carelli, King, Hampson, & Deadwyler, 1993; Chang, Sawyer, Lee, & Woodward, 1994; Peoples, Gee, Bibi, & West, 1998; Wise, 1996; Wise, 2002). Other forms of pathological behavior including shopping addiction, hypersexuality, and compulsive eating have been found to have roots in the dopaminergic system triggering chemical reactions in the brain comparable to substances of abuse (Comings et al., 1996; Lim, Ha, Choi, Kang, & Shin, 2012). It comes as no surprise, then, than the same would be observed with exercise -an activity that can induce a euphoric psychological state (Harber & Sutton, 1984). The pleasurable effects, behaviors, hormone levels, and potential for addiction experienced by individuals who exercise are similar to those elicited by other natural and non-essential rewards (Leuenberger, 2006).

Physical activity can be a self-rewarding behavior that can exhibit addictive properties (Kolb, Kelly, & Garland Jr., 2013) and significant departures from healthy activity levels have been associated with behavioral disorders (Majdak et al., 2014). While addiction to exercise is not considered an official addictive disorder (American Psychiatric Association, 2013), 3% to 5% of the US population experience this condition (Sussman, Lisha, & Griffiths, 2011). Exercise dependence involves the brain’s reward pathway and is attributed to endorphin hormones which in high levels act in a way similar to morphine and have been found to correlate with vigorous exercise (Goldfarb, Hatfield, Sforzo, & Flynn, 1987; McKim, 2003; Pierce, Eastman, Tripathi, Olson, &

Dewey, 1993). Physical activity causes release of dopamine in the basal ganglia (Pieau & Dorizzi, 2004), suggesting that it activates the dopaminergic mesocorticolimbic brain system. Also, the neural processes of physical activity overlap with those of rewards substances previously discussed. Likewise, withdrawal from physical activity has been associated with depression and anxiety behaviors (Kolb et al., 2013).

The psychological and biological characteristics shared between exercise dependence and reward-dependent behaviors have led to an effective model that uses mouse lines selected for higher physical activity compared to control unselected lines (Garland et al., 2010; Uyttenhove et al., 2003). A number of studies that used this model have been performed on the striatum because of its involvement in decision-making, goal-directed action (Balleine, Delgado, & Hikosaka, 2007) and association with motor behavior and exercise (Alexander, DeLong, & Strick, 1986; Penney & Young, 1983). Furthermore, exercise increases the concentration of dopamine in this brain region (De Bruin, Schasfoort, Steffens, & Korf, 1990; Kuhr & Korf, 1988a; Kuhr & Korf, 1988b; Schasfoort, De Bruin, & Korf, 1988; Teixeira et al., 2009; G. J. Wang et al., 2000). However, recent results from studies on mice selectively bred for high levels of voluntary wheel running indicate that in these hyperactive mice, physical activity may affect the cerebellum in a curious and unique way. While mice from these high-runner lines showed statistically larger non-cerebellar brain mass, their cerebellums did not differ from those of control mice (Kolb et al., 2013). Given that the cerebellum is linked to exercise but was unaltered by this selective breeding, it appears that voluntary exercise in highly active may have a distinctive impact on the cerebellum.

The cerebellum and striatum are closely linked. Studies have demonstrated that the output stages of cerebellar processing have a disynaptic connection with the input stages of basal ganglia processing, which occur in the striatum. (Hosni et al., 2005) However, although these two regions of the brain work together closely, their neurons have different targets and functions; additionally, they appear to implement different learning algorithms, which would impact physiological responses to reward (Doya, 2000; Houk, 2005). Supplementing studies on exercise and reward in the striatum with research performed on the cerebellum may be extremely beneficial for better understanding how various parts of the brain respond to reward, addiction, and exercise.

The goals of this study are to uncover changes in the cerebellum associated with genetic selection for physical activity-dependence and the modulation of these genetic changes associated with physical activity environment. We profile the cerebellum transcriptome of mice on a 2-by-2 factorial design comparing a line selected for high voluntary physical activity versus a parallel control line both under restricted and unrestricted running environmental conditions. The comparison of both genetics lines under restricted and unrestricted activity environments permitted to disentangle the transcriptome profiles and networks underlying activity dependence from the changes in transcriptome in response to coincidental physical activity. Gene networks including information on the transcriptome profiles characterized in this study were compared across genetic and environmental activity levels. Transcriptomic distinctiveness and similarities between genetic lines within environments and between environments within genetic lines will advance the understanding of genetic variation in reward dependency and addictive behaviors.

Materials and methods

Experimental Design

An experiment following a randomized 2 x 2 factorial design including the factors of activity or exercise genetics and environment was carried out. The design included two mouse lines, one line selected for high voluntary physical activity and an unselected control line (hereby denoted activity genotype), and two cage environments enabling two activity levels blocked from wheel running or allowed to run (hereby denoted activity environment). Sample size was n=4/activity genetic-environment combination. Mice were weaned at 21 days of age, kept in temperature-controlled rooms (21 ± 1°C) on a 12 hour photo-period, and provided food and water ad libitum (Swallow, Carter, & Garland, 1998). All procedures that involved animals were approved by the University of California, Riverside, Institutional Animal Care and Use Committee.

Activity genotype groups. Two lines, activity genotype levels high and control or low (H and L, respectively) were studied. Mice corresponded to a line selectively bred

for high voluntary wheel-running (H Line 7) and one non-selected control line (L Line 1) from an ongoing selective breeding program beginning with a mix of 8 Hsd:ICR strains. Details on the successful protocol of selection for high activity have been documented previously (Dlugosz et al., 2013; Hannon et al., 2010; Talmadge, Acosta, & Garland Jr., 2014; Waters et al., 2013). Briefly, selected lines were bred for higher voluntary distance run on a wheel within the habitual cage. The randomly-bred control line was raised under the same conditions. Each generation, at approximately 6 to 8 weeks of age, individual mice are placed in cages with access to a Wahman-type running wheel (circumference= 1.12 m) for 6 days. The selection criterion is the average count of revolutions run on days 5 and 6 tracked by a computer-automated system. High activity line parents of the next generation are the female and male from each litter that have the highest revolutions count meanwhile control line parents are two randomly chosen males and two females from each litter. Beyond the 6-day period when revolution counts are recorded, mice are maintained in randomly assigned, same-sex groups of four in standard cages until the environment treatment is evaluated. Male adult mice corresponding to approximately the 40^{th} generation of the selection experiment were used in this study. By this generation, the 4 high activity lines of mice run nearly 3 times more distance per day and at higher velocity relative to the four control lines (Swallow et al., 1998). The difference between selected and control groups is higher than the difference within group. The H and L lines used in this study were selected at random among the selected and control lines available.

High voluntary wheel running has been recognized a recognized self-rewarding behavior in mice (Kolb et al., 2013). The higher running of the selected lines has been associated with changes in dopamine and endocannabinoid signaling (Keeney, Meek, Middleton, Holness, & Garland, 2012; Rhodes, Gammie, & Garland, 2005). Furthermore, the high activity lines in these studies also exhibited high home-cage activity in the absence of wheel and high behavioral despair after removal of wheel following days of access (Kolb et al., 2013). The interpretation of the biological significance from the functional enrichment performed on the differentially expressed genes between activity genotype-environment combinations, genotypes and environments, it is valuable to understand previously identified phenotypic and

genotypic characteristics of the mice lines bred for high wheel running. The selectively bred mice display increased physical activity compared to control mice (voluntary wheel running and climbing) regardless of manipulation of external factors, such as maintaining them under constant dark or constant light (Koteja, Swallow, Carter, & Garland, 2003). They run faster, bouts of running occur more frequently, and they exit the wheel less frequently (Rhodes et al., 2005). Even when the running wheels are locked, mice from the selected lines climb more in their cages, and increased wheel rotational resistance does not discourage them from wheel-running (Koteja, Garland, Sax, Swallow, & Carter, 1999). Interestingly, while both control and selected mice were able to be trained to press a lever for a reward of 30 minutes of free running time, only control lines could be trained to perform the behavior for only 90 seconds of running time (Belke & Garland, 2007). These behaviors suggest a general higher motivation towards physical activity, and specifically longer running durations, compared to control mice. Physical features of mice selected for high wheel running include larger hearts, thicker femurs and tibiafibulas, larger femoral heads, and more symmetrical hindlimb bone lengths (Garland & Freeman, 2005). Mice selected for high wheel running have less body mass and length, less body fat even when housed without access to a wheel, and eat more for their body mass (Nehrenberg, Hua, Estrada-Smith, Garland, & Pomp, 2009).After accounting for their reduced body fat, they have lower circulating leptin levels and high circulating adiponectin and corticosterone levels, even at rest during the day (Majdak et al., 2014).

Activity environment groups. Two environments, activity unrestricted and restricted (activity environment levels H and L, respectively) were studied. These environments represent availability and unavailability of the activity reward. These environments were evaluated on both genotypes because the comparison of activity genotypes alone could lead to confounding between cause and effect (do transcript profiles spur or are a result of voluntary activity?). To address this, the comparison between activity environments enabled the identification of differential gene expression between activity restricted and unrestricted mice due to the ability to exercise rather than due to genotypic differences in voluntary activity. The activity restricted and unrestricted environments were previously described in detail (Kelly et al., 2014).

Briefly, at approximately 11 weeks of age mice were placed in individual cages and either allowed (unrestricted) or prevented (restricted) access to a wheel; for mice in the restricted environment, the wheel was locked using a wire tie to prevent it from rotating. Groups were labeled by genotype followed by environment. Therefore, the mice group HL correspond to mice with high activity genotype in a low activity (restricted) environment.

RNA profiling

Mice were euthanized by rapid decapitation 1 day after individual housing in running or sedentary environments. Brain was extracted and the whole cerebellum was dissected on a chilled aluminum block, flash-frozen in liquid nitrogen and stored at -80°C. Subsequently, samples were thawed, submerged in RNAlater solution (Qiagen, Valencia, CA), and maintained at -20°C until RNA isolation. Individual cerebellum samples were homogenized and RNA purified (RNeasy Mini kit, Qiagen, Valencia, CA), then quantified and assessed for purity using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE) and an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara CA). RNA integrity was assessed using the Agilent 2100 Bioanalyzer with RNA Pico chip (Agilent Technologies, Palo Alto, CA). All samples had RNA Integrity Numbers (RIN) > 9. Libraries of individual mouse cerebellum samples were sequenced (Nixon et al., 2015) and 100nt-long paired-end reads were obtained using Illumina HiSeq 2000 (Illumina, San Diego, CA, USA).

Analysis

Using the FastqGroomer tool, FastQ data files were converted to FastqSanger format following proven protocols (Caetano-Anollés, Mishra, & Rodriguez-Zas, 2015; Nixon et al., 2015). A quality control check was performed on the reads using FastQC, and end positions with low Phred quality values (<20) were removed with Fastq Quality Trimmer. Next, the reads were mapped to the mouse mm10 genome assembly (UCSC; http://genome.ucsc.edu) using Tophat (v2.0.13) (Trapnell et al., 2012). Tophat settings were expected (mean) inner distance = 200nt, anchor length" junctions spanned by reads with at least 8bp on each side of the junction with 0 mismatch, independent mapping of

read segments of 25nt long, and default intron length specifications. Cufflinks was used to assemble transcripts and estimates abundances; transcript isoforms were identified and quantified in number of fragments per kilobase of exon per million mapped reads. Differential transcript abundance between experimental conditions was tested using Cuffdiff (v2.2.0) using a geometric library normalization (Trapnell et al., 2012).

Beyond standard identification of differentially expressed genes between activity genotypes and environments, this study aimed to uncover complex interplay between these factors. Differential expression was tested between genotype-environment groups, between genotype groups and between environment groups. Evaluation of the interaction between genotypes (H or L) and environments (H or L) allowed the identification environment-dependent differential profiles between high activity and control genotypes (Caetano-Anollés et al., 2015). Differences between genotypes within environment offers insights into transcripts associated into behavioral differences between genotypes in presence or absence of the reward (wheel activity). These tests allowed assessment of the statistical significance of the interaction between genotype and environment on the transcriptome. Gene profiles exhibiting significant interaction effects were further evaluated for possible synergistic or antagonistic effects. Example of synergistic effect would be H genotype and H environment resulting on gene expression level more extreme than H genotype and H environment added. Example of antagonistic effect would be H genotype and H environment resulting on gene expression level less extreme than H genotype and H environment added. Testing for the main effects of genotype and environment followed the testing for interaction. False discovery rate (FDR) adjusted P-values were used to account for multiple test adjustment across transcripts.

Six pairwise contrasts were used to profile the expression patterns involving 4 groups: high activity genotype-high activity environment (HH); high activity genotype-low activity environment (HL); low activity genotype-high activity environment (LH); and low activity genotype-low activity environment (LL). Among the 6 pairwise contrasts, 3 orthogonal contrasts were of interest: 1) HH vs. LH, 2) HL vs. LL, and 3) HH vs. HL. The first two contrasts allowed the identification of differences between H and L genotypes within environment and the third contrasts allowed to identify

differences between environments within the H genotype. Commonalities between the first two contrasts support the identification of genes associated with high voluntary running regardless of environment. Commonalities between the first two groups and the third group support the identification of genes that may respond also to the environment.

Enrichment of functional terms was investigated using the differential gene expression from the three orthogonal contrasts of interest: 1) HH vs. LH, 2) HL vs. LL, and 3) HH vs. HL. Two complementary functional enrichment analyses of the transcriptome profiles were undertaken. First, among genes differentially expressed (FDR-adjusted P-value < .05 comparable to P-value < 0.001) enrichment of functional categories and pathways was performed using the Database for Analysis, Validation, and Integrated Discovery system (DAVID; (Huang da, Sherman, & Lempicki, 2009a; Huang da, Sherman, & Lempicki, 2009b)). Gene Ontology (GO) functional categories investigated included biological process (BP), molecular function (MF), and pathways defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG). GO Functional Annotation Tool (FAT) categories, which filter broad and less informative terms, along with clustering of categories that share genes were used to minimize redundancy between FAT and KEGG categories (Delfino & Rodriguez-Zas, 2013; Serao, Gonzalez-Pena, Beever, Bollero et al., 2013; Serao, Gonzalez-Pena, Beever, Faulkner et al., 2013). Each cluster's statistical significance was assessed using enrichment scores produced in DAVID; these values correspond to the geometric mean of the scores of each functional category in the cluster. Second, Gene Set Enrichment Analysis (GSEA) was performed on the gene expression profiles of all genes studied to uncover enriched categories within the up- and down-regulated genes between genotype-environment groups (Subramanian et al., 2005). Cuffdiff output files were converted to .gct format using the Fpkm_trackingToGct tool within Genepattern software (Subramanian et al., 2005) and then submitted to GSEA for analysis. Maximum and minimum cutoffs for gene set size were 1000 and 5 respectively, and enrichment significance was assessed using 1000 permutations. GSEA analyses confirmed results of the gene enrichment analysis performed using DAVID, which are interpreted subsequently.

Gene networks were reconstructed using the differential gene expression from the three orthogonal contrasts of interest: 1) HH vs. LH, 2) HL vs. LL, and 3) HH vs.

HL. From our list of 430 genes exhibiting significant activity genotype-by-environment interaction effect on gene expression at FDR P-value < 0.05, 99 were selected to visualize gene networks using the Cytoscape software (Killcoyne, Carter, Smith, & Boyle, 2009). In order to select these 99 genes, the 33 genes with the highest FDR adjusted P-value from each of the three orthogonal contrasts were selected and the relationships between these genes was visualized using the Bisogenet package within Cytoscape (Martin et al., 2010). Bisogenet was prompted to search for protein-protein interactions; the package utilizes data on interactions stored in publicly available data sets including BIOGRID, HPRD, DIP, BIND, INTACT, and MINT among others (Martin et al., 2010). All available data sources within BisoGenet were selected in order to generate the interactions.

Results and Discussion

Summary of RNA-Seq data

The quality and quantity of the RNA sequence reads was evaluated across samples. The average quality score Phred of the reads was 30. These scores are excellent, considering that a Phred of 30 is the equivalent of P-value < .001 (Z. Wang, Gerstein, & Snyder, 2009). The number of reads and quality scores along the reads were comparable across samples from all 6 activity genotype-environment groups. Likewise, the percentage of reads mapped to the mouse genome was 100% (an average of 12,9718676.5 total reads were mapped per sample). From these findings, genes corresponding to transcripts exhibiting statistical significant (FDR-adjusted P-value < 0.005) interaction between genotype and environment, or significant main effects of genotype or environment (FDR-adjusted P-value < 0.005; log2(fold change) > |1.3|) were identified and their profile characterized.

Activity genotype-by-environment interaction effect on gene expression

Overall, 430 genes exhibited significant interaction effect at FDR < 0.05. The top 20 genes exhibiting the most significant interaction effect including the log2(fold

change) between pairs of genotype-by-environment groups and the overall FDR-adjusted P-value are presented in Table 3.1.

Among the gene profiles exhibiting significant genotype-by-environment interaction effect, a number of genes have been previously associated with reward-processing and reward-dependent behaviors. Adenosine A2a receptor (Adora2) was over-expressed in the LH-HH contrast and under-expressed in all other contrasts. A2ar has been associated with the physiological response of the body to drug-use and hyperactivity (R. M. Brown & Short, 2008). This receptor is involved in goal-focused behaviors and motor functions used to carry out those behaviors and when this receptor is deleted, mice display reduced exploratory behaviors and even less motivation towards sexual activity and eating (Short, Ledent, Drago, & Lawrence, 2006). In rats, subchronic doses of morphine increases levels of extracellular adenosine and therefore decreases striatal A2a receptor number and function (De Montis, Devoto, Meloni, Saba, & Tagliamonte, 1992). Also, A2ar agonists have been found to significantly decrease cocaine induced hyperactivity in mice (Poleszak & Malec, 2002). These receptor is also tied to locomotor activity in dopamine D2 receptor (Drd2) deficient mice, locomotor impairment was reversed using an A2ar antagonist (Aoyama, Kase, & Borrelli, 2000).

Consistent with the A2ar finding, our results revealed significant differential expression of Drd2. In our study, Drd2 was over-expressed in the LH-HH and HL-HH contrasts and under-expressed in all other contrasts. Dopamine plays a major role in inducing behavioral and physiological changes. Exercise increases levels of serum calcium, which is transported to the brain and subsequently stimulates dopamine synthesis (Sutoo & Akiyama, 2003). In mice lacking D2 receptors, the rewarding property of morphine was reduced although the behavioral expression or drug withdrawal was unchanged (Maldonado et al., 1997). This indicates that the Drd2 is crucial for the motivational component of drug addiction, although it may have a lesser effect on the physiological response of the body to addiction. These results suggest that both drug abuse and hyperactivity (in the selection line studied) are reward-motivated behaviors that involve similar molecular mechanisms in the brain.

Comparison of gene expression across activity genotype-by-environment groups

The number of differentially expressed genes at FDR adjusted P-value < 0.05 (and corresponding contrast) was: for 117 (LL-HL), 140 (LL-LH), 121 (LL-HH), 261 (HL-LH), 12 (HL-HH), and 252 (LH-HH). Within these contrasts, the number of over-expressed genes was: 12, 4, 8, 100, 5, and 130 respectively. Several interpretations can be made from the progression of the number of differentially expressed genes starting with the highest number in LH-HL followed by LH-HH, followed by LL-LH followed by LL-HH. Firstly, the number of differentially expressed genes between genotypes regardless of environment (LL-HL and LH-HH) was higher than the number of differentially expressed genes between environments regardless of genotype (LL-LH and HL-HH). This finding indicates that the impact of genotype is more prevalent than that of environment on the transcriptome. Also, genotype differences were associated with more than double the number of differentially expressed genes in H environment (LH-HH) relative to L environment (LL-HL). This result may suggest that the genotype differences in gene expression may benefit from an environment that supports high activity. **Figure 3.1** presents a Venn diagram illustrating the shared differentially expressed genes between on HH vs LH, HL vs LL, and HH vs HL; observing these contrasts highlight genetic differences and those which may be related to environment. The limited overlap between contrasts depicted in this diagram highlights that the differences in gene expression between genotypes are highly dependent on the environment. While there were no genes that were differentially expressed in the three contrasts, a substantial number of genes were expressed in both HL-LL and HH-LH (16), followed by HH-LH and HH-HL (6) and HL-LL and HH-HL (2). Sharing indicates trigger of similar molecular mechanisms across environments.

Another finding from the comparison of differentially expressed genes across contrasts is that environmental differences were associated with significantly more differentially expressed genes (almost tenfold) in L genotype (LL-LH, 116 genes) than in the H genotype (HL-HH, 11 genes). This result indicates that environment can elicit differential gene expression in the L genotype that may already be activated in the H genotype and thus the limited differential expression response to the environment.

Among all genotype-environment contrasts, HL-LH mice exhibited the highest number of differentially expressed genes relative to other genotype-environments

contrasts. We conclude that the changes in transcriptome associated with activity genotype are not balanced out by changes associated with activity environment. The remaining contrast involving different genotypes and environment (LL-HH) exhibited half of the number of differentially expressed genes than the contrasts with most differentially expressed genes (HL-LH and LH-HH). These findings could imply that there may be synergistic effects between opposite levels of genotype and environment resulting in the highest number of gene differences and antagonistic effects between similar levels of genotype and environment that cancel some of the gene expression differences.

Figure 3.2 illustrates the shared differentially expressed genes between one set of orthogonal contrasts including the LL baseline group. Unlike in the previous comparison of three contrasts, HH vs LH, HL vs LL, and HH vs HL, where each contrast appeared to have a unique set of differential genes with none shared between all 3, the comparison of all 3 contrasts including the LL baseline group showed a very different pattern. Overall, 81 genes were shared between all 3 contrasts, strongly suggesting a common molecular mechanisms operating across environments. This complementary view confirms that genotypic make-up leading to voluntary high activity cause changes in gene expression that environment alone cannot replicate.

Enriched functional categories

Functional enrichment analysis was performed individually on the 3 orthogonal contrasts between genotype-environment combinations of interest. Enriched DAVID clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes differentially expressed (FDR-adjusted P-value < 0.05) between the contrasts are provided in Tables 3.2 and 3.3.

Table 3.2 lists enriched (enrichment score > 3.0) clusters among the genes significantly differentially expressed between genotypes within the L environment (HL-LL). A central theme is neuron morphogenesis and the significant terms included: neuron and neuron projection, differentiation, cell and cellular component morphogenesis and morphogenesis involved in differentiation, axonogenesis and axon

guidance, regionalization. The terms “forebrain and telencephalon development”, “forebrain and telencephalon cell migration” together with a whole suite of ‘developments’ (cortex, hippocampus, diencephalon) are all related to neurogenesis and further supported by "cerebral cortex GABAergic interneuron".

Table 3.3 lists enriched (enrichment score > 3.0) clusters among the genes significantly differentially expressed between genotypes within the H environment (HH-LH). Enriched clusters included terms related to cell development, neurogenesis, nervous system development, and homeostatic processes. The former terms related to neurogenesis are consistently with enriched terms in the HL-LL contrasts. This finding suggests a main effect of high voluntary running genotype independent of reward availability (e.g. restricted or enabled wheel running). This is consistent with studies reporting association between physical activity and neural development (Kim et al., 2004; Noonan, Bulin, Fuller, & Eisch, 2010). However, our finding cannot discriminate whether changes in processes associated with neural development lead to changes in voluntary physical activity or the selection for voluntary physical activity lead to changes in neurological development. The identification of significant terms related to hormone metabolism, including cellular hormone metabolic process, regulation of hormone levels, and hormone metabolic process is unique to the HL-LL contrast and suggests an interaction between activity genotype and environment. This finding suggests that reward restriction elicits changes in genes corresponding to hormonal metabolic pathways. Hormone activity influences an individual’s behavior as well as the makeup of the brain (Arnold & Breedlove, 1985; Arnold, 2009) Testosterone levels are well known to impact the extent of novelty-seeking and sensation seeking behaviors (Ehrenkranz, Bliss, & Sheard, 1974; Soler, Vinayak, & Quadagno, 2000), and androgens and estrogens also elicit a dopamine response in the mesolimbic pathway (de Souza Silva, Mattern, Topic, Buddenberg, & Huston, 2009; Lammers et al., 1999; Landry, Levesque, & Di Paolo, 2002; McEwen, 2002). This enriched group is composed of 4 genes- Aldehyde Dehydrogenase 1 Family Member A1 (Alsh1a1), Lecithin Retinol Acyltransferase (Lrat), Cytochrome P450 11A1 (Cyp11a1), VGF Nerve Growth Factor Inducible (Vgf). Our findings offer valuable leads into hormone metabolism pathways role in reward

dependent behavior (e.g. high voluntary activity) on reward-absent environment (e.g. restricted activity).

Only 1 functional cluster among genes significantly (FDR-adjusted P-value < 0.05) differentially expressed between genotypes within the HH-HL contrast was identified and had a low enrichment score of 0.1. This cluster included three terms-metal ion binding, cation binding, and ion binding. As ion binding plays an important role in muscle function and plasticity (Berchtold, Brinkmeier, & Muntener, 2000), the presence of these terms in this exercise contrast is consistent what is currently known of the biochemical changes associated with muscle contraction and physical activity (Davis, Donovitz, & Hood, 1992; Fajardo et al., 2013; Kirber, Walsh, & Singer, 1988).

DAVID functional enrichment analysis results for HH-LH, HL-LL, and HH-HL were confirmed in GSEA. Unlike DAVID, GSEA can show whether functionally enriched clusters are up- or down-regulated. The "Hormone Metabolic Process" term was found significantly under-expressed in all three contrasts, and "Neurogenesis" was found over-expressed in results of analyses performed on HL-LL and HH-HL groups. Interestingly, this term was found significantly under-expressed in HH-LH.

Gene network analysis

Visualization of networks of genes differentially expressed in the three orthogonal contrasts of interest augmented our understanding of the molecular relationships impacted by the interaction between activity genotype and environment and by the corresponding main effects. Figures **3.3** and 3.**4** depict the networks of differentially expressed genes (FDR P-value < 0.05) between activity genotypes within activity environment (HL-LL and HH-LH, respectively). **Figure 3.5** depicts the network of differentially expressed genes (FDR P-value < 0.05) between activity environments within high voluntary activity (HL-HH). Nodes represent genes and edges represent the known associations between genes. The highest under-expressed genes with highest fold change were located in sparsely connected peripheral positions of the networks in contrasts between activity genotypes within activity environment (Figures 3.3 and 3.4). Lower level under expression occurred in hubs of highly connected genes. This is in

sharp contrast to the moderate overexpression of those same hubs genes in the contrast between activity environments within high voluntary activity (Figure 3.5).

Four genes exhibit a similar differential pattern in the HL-LL and HH-LH gene networks. Laminin, Gamma 2 (Lamc2), Fas Cell Surface Death Receptor (Fas), Ribosomal Protein L35a (Rpl35a), and LIM Homeobox 9 (Lhx9) are significantly over-expressed genes in the high voluntary running line relative to Control in both activity environments. The association of these genes with activity dependence, irrespectively of reward availability suggest a fundamental disruption that the environment is not able to alter. Lxhr9 regulates orexin, a neuropeptide involved in arousal, appetite, wakefulness, and reward processes (Liu et al., 2015). The involvement of orexin in the mechanisms of reward systems has led to suggestions that this gene could be linked to drug addiction (Tsujino & Sakurai, 2009). The other 3 genes also have been associated with addictive behaviors. Lamc2 is down-regulated by acute nicotine exposure (J. Wang et al., 2011). Fas receptor proteins are modulated during opiate addiction in the rat brain (GarcÃ-a-Fuster, Ferrer-AlcÃ³n, Miralles, & GarcÃ-a-Sevilla, 2003).

On the other hand, 6 genes- Neurogranin (Nrgn), dopamine receptor D2 (Drd2), Retinoid X Receptor Gamma (Rxrg), Guanine Deaminase (Gda), Adenosine A2a Receptor (Adora2a), and Ras-Related Protein Rab-40B (Rab40b)- display strikingly dissimilar patterns between the HL-LL and HH-LH gene networks. These genes are over-expressed in HH-LH and under-expressed in HL-LL. These genes that have different in expression patterns across environment exhibit interaction effect, where the effect of activity genotype depends on the reward availability environment. Nrgn is a neuropeptide associated with alcohol addiction through the role in synaptic plasticity and signal transduction (Levran et al., 2015). Drd2 has been associated with addiction to opiods, nicotine, and cocaine (Noble, 2003); however our study is first on uncovering the dependency of the profile on activity environment. The dramatic expression difference indicates that the relationship between Drd2 and addictive behaviors in mice motivated for physical activity as the reward has different mechanisms than those of addiction to drugs of abuse.

Finally, genes that were not highly differentially expressed (intermediate colors) in the two activity genotype contrasts included histone cluster 2, H4 (Hist2h4),

Neurogenin 2 (Neurog2), Ak181631, and minichromosome maintenance complex component 6 (Mcm6). These results indicate that high voluntary activity is not clearly associated with dysregulation of these genes and consequently these genes do not play a strong role in physical activity addiction. However, some of these genes have been associated with other types of addiction. Neurog2 is required for midbrain DA neuron development in mice (Andersson, Jensen, Parmar, Guillemot, & Bjorklund, 2006; Kele et al., 2006) and dysfunction of midbrain DA neurons is linked to drug addiction (Lupica & Riegel, 2005; Park et al., 2008).

Comparison between the networks highlighting genes differentially expressed between activity genotypes and the network highlighting genes differentially expressed between activity environments within the high activity genotype group offers insights into gene dysregulation associated with genotypes that can be mimicked or obliterated by the environment. Interestingly, most of the genes over- or under- expressed in the activity genotype comparison are only intermediately differentially expressed between activity environments. This result suggest that environment is not confounding our previous conclusions based on activity genotype comparisons. The only exception is Midline 1 (Mid1). This gene was significantly over-expressed in the HL-LL contrast yet was significantly under-expressed in the HH-LH. In fact, Mid1 was the only significantly under-expressed in this network. This finding suggests that reward availability environment can counteract the dysregulation in Mid1 expression associated with high voluntary activity genotype.

Effect of activity genotype on gene expression in the cerebellum

Table 3.4 summarizes genes significantly differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |3|) between low and high activity genotypes. Results from the functional enrichment analysis of the genes differentially expressed between activity genotypes (P-value < 0.001), and enriched (enrichment score > 3.0) clusters are highlighted in Table 3.5.

Dopamine receptor D1 (Drd1a) was overexpressed in H relative to L genotypes. Our results are consistent with reports that this receptor is key in mediating the effects of

drugs and has a strong role in reward-dependent behaviors (Tobon, Catuzzi, Cote, Sonaike, & Kuzhikandathil, 2015). Both D1-like and D2-like dopamine receptors (two classes of dopamine receptors; regulated by Drd1a and Drd2, respectively) are involved in rats who are taught to perform a behavior in order to receive pleasurable electrical stimulation (Franklin & Vaccarino, 1983; Kornetsky & Esposito, 1981). In mice addicted to cocaine, D1-like receptors regulated by Drd1a act to reduce the desire for further cocaine reinforcement (Self, Barnhart, Lehman, & Nestler, 1996). This result indicates that D1-like receptor agonists may be an attractive option for cocaine addiction and could be of value in understanding addiction to high voluntary activity. The role of the D1 receptor has been explored in studies of drug self-administration. Both D1 and D2 receptors are involved in the reward response and addiction to drugs of abuse; endogenous DA simulates D1 receptors, which leads to the expression of D2 receptor-mediated reward-driven behaviors and gene expression changes (Beninger, Hoffman, & Mazurski, 1989; Maldonado, Robledo, Chover, Caine, & Koob, 1993; Phillips, Robbins, & Everitt, 1994; Self et al., 1996). As this gene was found overexpressed in our study, it appears that this gene's process and function may be similar in mice seeking an exercise reward as those seeking drugs of abuse.

Retinoid X receptor gamma (Rxrg) was over-expressed in H relative to L activity genotype. Retinoid receptors directly up-regulate Drd2 (Samad, Krezel, Chambon, & Borrelli, 1997). Consistent with our findings, this gene has been found to directly impact ambulatory activity and Rxrg knock-out mice show impaired locomotion along with reduced dopamine signaling (Krezel et al., 1998). Also, ablation/loss of Rxrg signaling in mice leads to depressive behaviors due to the decreased dopamine signaling in the striatum, supporting the association of this gene with reward and pleasure (Krzyzosiak et al., 2010). Over-expression of Rxrg in humans is associated with sensation seeking (Alliey-Rodriguez et al., 2011). This is consistent with our finding of the over-expression of this gene in this study; just as this gene impacts sensation seeking in humans addicted to psychoactive substances, this gene may similarly impact the drive of selected mice to perform physical activity with pleasure as the sole incentive.

Muscarinic Acetylcholine Receptor M1 (Chrm1) was over-expressed in H relative to L activity genotype. The differentially expression of yet another gene related

to dopamine transmission (Nathanson, 1987; Wess, 1996) further confirms that activity, drug and other reward-dependent behaviors share disruption of dopamine pathways. In agreement with our findings, Chrm1 deficient mice display elevated dopamine transmission in the striatum (Gerber et al., 2001).

Effect of activity environment on gene expression in the cerebellum

Table 3.6 summarizes genes significantly differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |3|) between control and L environments. Results from the functional enrichment analysis of the genes differentially expressed between activity environments (P-value < 0.001) and enriched (enrichment score > 3.0) clusters are highlighted in Table 3.7.

Functional analysis of the genes differentially expressed between activity environments identified the enrichment of voltage gated channel activity and ion binding terms. Among the genes affiliated to these terms are those that code for transmembrane proteins that move cations and anions through membranes with energy costs. G-protein coupled receptors, which are involved in the energetics of these processes, are also enriched among the genes differentially expressed between activity environments (Kroeze, Sheffler, & Roth, 2003). Enrichment in this contrast group appears to be more related instant physiological responses while enrichment of genes differentially expressed between genotype and genotype-by-environment groups included the effects of selection appear to be related to cellular development and long-term responses occurring in the brain.

CONCLUSIONS

Exercise dependence and addiction to psychoactive substances share psychologically and physiological characteristics. A transcriptome study of a mouse line selected for high voluntary activity and compared to a control line in environments restricting or enabling rewarding activity was undertaken. Results offered insights into genetic factors behind activity dependence and provide a model for understanding addiction and other reward-dependent behaviors.

Gene network analysis identified connected genes that exhibited similar (e.g. Lhx9) and opposite (e.g. Nrgn) expression patterns between activity genotypes across reward availability environment. Meanwhile the former set of genes represent fundamental dysregulation, irrespectively of reward availability, the dysregulation of the second set of genes due to activity genotype can be altered by the reward availability. Most of the genes over- or under- expressed in the activity genotype comparison were only intermediately differentially expressed between activity environments, suggesting that environment did not confound our previous conclusions based on activity genotype comparisons.

Adora2a had a significant genotype-by-environment interaction effect evidenced by over-expression in the genotype contrast within high activity environment and under-expressed in other contrasts. This gene is involved in goal-focused behaviors and motor functions used to carry out those behaviors and our finding suggest that for activity addiction, the effect is dependent on the availability of the reward. A central theme from the functional analysis of activity genotype-environment was neuron morphogenesis. The enrichment of this and related biological processes is consistent with previous reports of association between exercise and neurological development.

Our findings of differentially expressed genes related to dopaminergic transmission are consistent with previous reports of increased locomotor activity correlated with elevated levels of dopamine in rodents. Drd1a was found to be overexpressed in the high voluntary activity relative to the control genotype. Furthermore, the Drd2, was also found overexpressed in high voluntary activity relative to the control genotype and is primarily involved in the motivational component of drug addiction [57]. Rxrg was over expressed in the high voluntary activity relative to the control genotype and is associated with higher ambulatory activity and sensation seeking in humans.

These results suggest that some of the molecular mechanisms underlying transcriptomic changes in mice selected for high voluntary activity are shared with other reward-dependent behaviors. Our findings support that high voluntary activity selection lines in mice are a helpful model to understand molecular mechanisms behind addition.

Also, identification of genes and biological processes associated with both high

voluntary activity and the pleasurable neurological response to physical activity may allow for the development of drugs and gene therapies which make it more pleasurable for people to exercise or less pleasurable to be sedentary as a treatment approach for overweight/obesity as well as a way to improve health overall in the general population.

Acknowledgements

The support of NIH grant numbers R21 DA027548 and P30 DA018310 and USDA NIFA grant numbers ILLU-538-632 and ILLU-538-909 is greatly appreciated.

TABLES AND FIGURES

Table 3.1 Genes exhibiting significant (FDR-adjusted P-value < 0.005, average log2(fold change) > |2|) activity genotype-by-environment interaction

Gene Name *	Log2Fold[1]						FDR-adjusted P-Value	P-Value	Average of all |Log2 Fold|
	LL - HL	LL - LH	LL-HH	HL - LH	HL-HH	LH-HH			
Ddn	-5.1239	-7.6548	-6.92317	-2.5309	-1.79927	0.731627	2.57307E-14	1.11022E-16	3.537666714
AK049668	2.36733	-3.81808	2.4553	-6.18542	0.087979	6.27339	2.57307E-14	1.11022E-16	3.027057553
Gpr88	-4.48439	-6.42635	-5.93863	-1.94196	-1.45425	0.487719	2.57307E-14	1.11022E-16	2.961899857
Gda	-4.60835	-6.54463	-5.29571	-1.93628	-0.687357	1.24892	2.57307E-14	1.11022E-16	2.903035286
Icam5	-4.57647	-6.21878	-5.70006	-1.64231	-1.12413	0.518182	2.57307E-14	1.11022E-16	2.825781714
Kcnj4	-4.596	-6.17144	-5.47734	-1.57544	-0.881339	0.694103	2.57307E-14	1.11022E-16	2.770808857
Arhgap8	2.06692	-3.50808	2.25669	-5.575	0.189776	5.76478	2.31366E-11	1.56875E-13	2.767909329
Scd4	3.28115	-1.66477	3.17214	-4.94592	-0.109003	4.83692	2.57307E-14	1.11022E-16	2.597429143

Table 3.1 (cont.)

Lrrc10b	-4.43931	-5.64017	-5.25873	-1.20086	-0.819418	0.38144	2.57307E-14	1.11022E-16	2.534275429
Cpne5	-4.54153	-5.76022	-4.817	-1.21869	-0.275474	0.943217	2.57307E-14	1.11022E-16	2.508018714
Nrgn	-4.50147	-5.50741	-4.02525	-1.00594	0.47622	1.48216	2.57307E-14	1.11022E-16	2.42835
Rxrg	-3.84299	-5.28096	-4.36536	-1.44661	-0.522372	0.924239	2.57307E-14	1.11022E-16	2.341595857
Cyp11a1	1.4743	-3.17024	1.57059	-4.64454	0.0962909	4.74083	8.97957E-06	1.50367E-07	2.250391514
Tbr1	-3.99168	-4.93814	-4.6618	-0.946458	-0.670122	0.276336	2.57307E-14	1.11022E-16	2.212076572
Adora2a	-4.02261	-4.8696	-4.72083	-0.846991	-0.698221	0.14877	2.57307E-14	1.11022E-16	2.186717429
Drd2	-4.18193	-4.83802	-4.09709	-0.65609	0.0848317	0.740922	2.57307E-14	1.11022E-16	2.085554815
Lamp5	-2.9593	-4.53335	-3.87565	-1.57405	-0.916353	0.657694	2.57307E-14	1.11022E-16	2.073771001
Sst	-4.16424	-4.60927	-3.54259	-0.445032	0.621649	1.06668	2.57307E-14	1.11022E-16	2.064208715
Cpne4	-3.45958	-4.77135	-3.56315	-1.31177	-0.103572	1.2082	2.57307E-14	1.11022E-16	2.059660292
Nts	-3.36256	-4.72986	-3.36298	-1.36731	-0.000425478	1.36688	5.21477E-08	5.57149E-10	2.027149268

1. Log2Fold is the log2-transformed fold change between two genotype-environment groups. When considering groups A vs B = A - B, Log2Fold Change = Log2 (A/B). Groups are: LL=low activity genotype-low activity environment; LH=low activity genotype-high activity environment; HL=High activity genotype-low activity environment; HH=High activity genotype-high activity environment

*Expanded gene names, listed in alphabetical order: 4933409K07Rik= RIKEN cDNA 4933409K07 gene; Adi1= acireductone dioxygenase 1; Adora2a= adenosine A2a receptor; AK049668= ; Cntn4= contactin 4; Cox7b= cytochrome c oxidase subunit VIIb; Cpne4= copine IV; Cpne5= copine V; Ctxn1= cortexin 1; D430019H16Rik= RIKEN cDNA D430019H16 gene; Ddn= dendrin; Dnah1= dynein, axonemal, heavy

chain 1; Drd2= dopamine receptor D2; Gda= guanine deaminase; Gm1976= predicted gene 1976; Gpr88= G-protein coupled receptor 88; Gsn= gelsolin; Icam5= intercellular adhesion molecule 5, telencephalin; Kcnj4= potassium inwardly-rectifying channel, subfamily J, member 4; Lamp5= lysosomal-associated membrane protein family, member 5

Table 3.2 Enriched (enrichment score > 2.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes significantly (P-value < 0.001) differentially expressed between Low activity Control genotype – Low activity environment vs High activity genotype – Low activity environment(LL-HL) mice.

Category	Term	Count	PValue	FDR
Score = 4.54				
GOTERM_BP_FAT	GO:0030900~forebrain development	11	1.91E-08	2.99E-05
GOTERM_BP_FAT	GO:0031175~neuron projection development	12	2.10E-08	3.27E-05
GOTERM_BP_FAT	GO:0030182~neuron differentiation	15	2.16E-08	3.37E-05
GOTERM_BP_FAT	GO:0048812~neuron projection morphogenesis	11	3.16E-08	4.94E-05
GOTERM_BP_FAT	GO:0048666~neuron development	13	4.42E-08	6.91E-05
GOTERM_BP_FAT	GO:0048858~cell projection morphogenesis	11	1.16E-07	1.81E-04
GOTERM_BP_FAT	GO:0032990~cell part morphogenesis	11	1.82E-07	2.85E-04
GOTERM_BP_FAT	GO:0007409~axonogenesis	10	2.08E-07	3.24E-04
GOTERM_BP_FAT	GO:0006928~cell motion	13	5.25E-07	8.20E-04
GOTERM_BP_FAT	GO:0048667~cell morphogenesis involved in neuron differentiation	10	5.29E-07	8.26E-04
GOTERM_BP_FAT	GO:0007411~axon guidance	8	9.03E-07	0.00141
GOTERM_BP_FAT	GO:0030030~cell projection organization	12	9.77E-07	0.001527
GOTERM_BP_FAT	GO:0000904~cell morphogenesis involved in differentiation	10	1.89E-06	0.002954
GOTERM_BP_FAT	GO:0000902~cell morphogenesis	11	5.59E-06	0.008736
GOTERM_BP_FAT	GO:0032989~cellular component morphogenesis	11	1.70E-05	0.026589
GOTERM_BP_FAT	GO:0021954~central nervous system neuron development	5	2.79E-05	0.04355
GOTERM_BP_FAT	GO:0021953~central nervous system neuron differentiation	5	7.35E-05	0.114815

Table 3.2 (cont.)

GOTERM_BP_FAT	GO:0007389~pattern specification process	7	0.00417	6.320529
GOTERM_BP_FAT	GO:0003002~regionalization	6	0.005835	8.73659
GOTERM_MF_FAT	GO:0003700~transcription factor activity	11	0.006469	7.736076
GOTERM_MF_FAT	GO:0043565~sequence-specific DNA binding	9	0.007916	9.389923
GOTERM_MF_FAT	GO:0030528~transcription regulator activity	12	0.045488	43.87611
GOTERM_BP_FAT	GO:0006355~regulation of transcription, DNA-dependent	14	0.046954	52.82693
GOTERM_BP_FAT	GO:0051252~regulation of RNA metabolic process	14	0.052108	56.6588
GOTERM_BP_FAT	GO:0045449~regulation of transcription	18	0.079545	72.60868
GOTERM_MF_FAT	GO:0003677~DNA binding	14	0.126552	81.33892
GOTERM_BP_FAT	GO:0006350~transcription	12	0.352403	99.88728
Score = 3.47				
Category	Term	Count	PValue	FDR
GOTERM_BP_FAT	GO:0030900~forebrain development	11	1.91E-08	2.99E-05
GOTERM_BP_FAT	GO:0021537~telencephalon development	7	2.30E-06	0.003598
GOTERM_BP_FAT	GO:0021892~cerebral cortex GABAergic interneuron differentiation	4	5.13E-06	0.00801
GOTERM_BP_FAT	GO:0021954~central nervous system neuron development	5	2.79E-05	0.04355
GOTERM_BP_FAT	GO:0021895~cerebral cortex neuron differentiation	4	3.16E-05	0.049375
GOTERM_BP_FAT	GO:0021953~central nervous system neuron differentiation	5	7.35E-05	0.114815
GOTERM_BP_FAT	GO:0021885~forebrain cell migration	4	1.36E-04	0.211512
GOTERM_BP_FAT	GO:0021830~interneuron migration from the subpallium to the cortex	3	1.70E-04	0.26468
GOTERM_BP_FAT	GO:0021853~cerebral cortex GABAergic interneuron migration	3	1.70E-04	0.26468

Table 3.2 (cont.)

GOTERM_BP_FAT	GO:0021894~cerebral cortex GABAergic interneuron development	3	1.70E-04	0.26468
GOTERM_BP_FAT	GO:0021543~pallium development	5	1.84E-04	0.287046
GOTERM_BP_FAT	GO:0021826~substrate-independent telencephalic tangential migration	3	2.82E-04	0.439238
GOTERM_BP_FAT	GO:0021843~substrate-independent telencephalic tangential interneuron migration	3	2.82E-04	0.439238
GOTERM_BP_FAT	GO:0016477~cell migration	7	0.001809	2.789374
GOTERM_BP_FAT	GO:0022029~telencephalon cell migration	3	0.00412	6.246199
GOTERM_BP_FAT	GO:0048870~cell motility	7	0.00417	6.320529
GOTERM_BP_FAT	GO:0051674~localization of cell	7	0.00417	6.320529
GOTERM_BP_FAT	GO:0021766~hippocampus development	3	0.007279	10.78636
GOTERM_BP_FAT	GO:0021761~limbic system development	3	0.013499	19.13005
GOTERM_BP_FAT	GO:0021987~cerebral cortex development	3	0.016794	23.24853
GOTERM_BP_FAT	GO:0021536~diencephalon development	3	0.017666	24.30521
GOTERM_BP_FAT	GO:0042493~response to drug	3	0.08113	73.33619
Score = 3.08				
Category	Term	Count	PValue	FDR
GOTERM_BP_FAT	GO:0030900~forebrain development	11	1.91E-08	2.99E-05
GOTERM_BP_FAT	GO:0060284~regulation of cell development	7	2.08E-04	0.324851
GOTERM_BP_FAT	GO:0050767~regulation of neurogenesis	6	7.10E-04	1.103177
GOTERM_BP_FAT	GO:0045665~negative regulation of neuron differentiation	4	7.90E-04	1.226933
GOTERM_BP_FAT	GO:0051960~regulation of nervous system development	6	0.001187	1.838294
GOTERM_BP_FAT	GO:0045596~negative regulation of cell differentiation	6	0.002938	4.492073

Table 3.2 (cont.)

GOTERM_MF_FAT	GO:0043565~sequence-specific DNA binding	9	0.007916	9.389923
GOTERM_BP_FAT	GO:0045664~regulation of neuron differentiation	4	0.01748	24.08129
GOTERM_BP_FAT	GO:0045165~cell fate commitment	3	0.186894	96.0533
Score = 3.02				
Category	Term	Count	PValue	FDR
GOTERM_BP_FAT	GO:0021954~central nervous system neuron development	5	2.79E-05	0.04355
GOTERM_BP_FAT	GO:0021953~central nervous system neuron differentiation	5	7.35E-05	0.114815
GOTERM_BP_FAT	GO:0050767~regulation of neurogenesis	6	7.10E-04	1.103177
GOTERM_BP_FAT	GO:0051960~regulation of nervous system development	6	0.001187	1.838294
GOTERM_BP_FAT	GO:0008284~positive regulation of cell proliferation	3	0.453056	99.99195
Score = 2.63				
Category	Term	Count	PValue	FDR
GOTERM_MF_FAT	GO:0022843~voltage-gated cation channel activity	7	5.18E-05	0.064279
GOTERM_MF_FAT	GO:0005249~voltage-gated potassium channel activity	6	1.60E-04	0.198065
GOTERM_MF_FAT	GO:0022832~voltage-gated channel activity	7	3.26E-04	0.403855
GOTERM_MF_FAT	GO:0005244~voltage-gated ion channel activity	7	3.26E-04	0.403855
GOTERM_MF_FAT	GO:0005267~potassium channel activity	6	5.45E-04	0.674557
GOTERM_BP_FAT	GO:0015672~monovalent inorganic cation transport	8	0.001184	1.833755
GOTERM_BP_FAT	GO:0006813~potassium ion transport	6	0.001676	2.5872
GOTERM_BP_FAT	GO:0006812~cation transport	10	0.001688	2.605178
GOTERM_MF_FAT	GO:0005261~cation channel activity	7	0.001786	2.193625

Table 3.2 (cont.)

GOTERM_BP_FAT	GO:0030001~metal ion transport	9	0.002485	3.812287
GOTERM_MF_FAT	GO:0030955~potassium ion binding	5	0.003276	3.989465
GOTERM_MF_FAT	GO:0022836~gated channel activity	7	0.003332	4.055913
GOTERM_MF_FAT	GO:0046873~metal ion transmembrane transporter activity	7	0.003888	4.718625
GOTERM_MF_FAT	GO:0031420~alkali metal ion binding	6	0.004293	5.197335
GOTERM_MF_FAT	GO:0005216~ion channel activity	7	0.009395	11.05167
GOTERM_MF_FAT	GO:0022838~substrate specific channel activity	7	0.010842	12.65057
GOTERM_MF_FAT	GO:0022803~passive transmembrane transporter activity	7	0.01155	13.4228
GOTERM_MF_FAT	GO:0015267~channel activity	7	0.01155	13.4228
GOTERM_BP_FAT	GO:0006811~ion transport	10	0.013677	19.35736
GOTERM_BP_FAT	GO:0055085~transmembrane transport	6	0.100989	81.04721
Score = 2.08				
Category	Term	Count	PValue	FDR
KEGG_PATHWAY	mmu04080:Neuroactive ligand-receptor interaction	8	4.75E-05	0.041743
GOTERM_BP_FAT	GO:0007186~G-protein coupled receptor protein signaling pathway	16	0.071843	68.80041
GOTERM_BP_FAT	GO:0007166~cell surface receptor linked signal transduction	18	0.172778	94.83569

Table 3.3 Enriched (enrichment score > 2.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes significantly (P-value < 0.001) differentially expressed between high activity genotype-high activity environment and low activity Control genotype- high activity environment (HH-LH) mice.

Category	Term	Count	PValue	FDR
Score = 1.20				
GOTERM_BP_FAT	GO:0042592~homeostatic process	11	0.016696	22.9335
GOTERM_BP_FAT	GO:0048878~chemical homeostasis	8	0.025178	32.60153
GOTERM_BP_FAT	GO:0030003~cellular cation homeostasis	5	0.032022	39.56209
GOTERM_BP_FAT	GO:0006875~cellular metal ion homeostasis	4	0.044328	50.41698
GOTERM_BP_FAT	GO:0055065~metal ion homeostasis	4	0.051153	55.62173
GOTERM_BP_FAT	GO:0006873~cellular ion homeostasis	6	0.055255	58.49857
GOTERM_BP_FAT	GO:0055080~cation homeostasis	5	0.057407	59.93774
GOTERM_BP_FAT	GO:0055082~cellular chemical homeostasis	6	0.060531	61.94312
GOTERM_BP_FAT	GO:0050801~ion homeostasis	6	0.081636	73.22354
GOTERM_BP_FAT	GO:0030005~cellular di-, tri-valent inorganic cation homeostasis	4	0.089087	76.3942

Table 3.3 (cont.)

GOTERM_BP_FAT	GO:0055066~di-, tri-valent inorganic cation homeostasis	4	0.10811	82.97044
GOTERM_BP_FAT	GO:0019725~cellular homeostasis	6	0.133966	89.19755
GOTERM_BP_FAT	GO:0006874~cellular calcium ion homeostasis	3	0.16085	93.36831
GOTERM_BP_FAT	GO:0055074~calcium ion homeostasis	3	0.171999	94.60801
Score = 1.04				
GOTERM_BP_FAT	GO:0060284~regulation of cell development	5	0.03686	44.07099
GOTERM_BP_FAT	GO:0050767~regulation of neurogenesis	4	0.086067	75.15385
GOTERM_BP_FAT	GO:0051960~regulation of nervous system development	4	0.111424	83.92359
GOTERM_BP_FAT	GO:0045664~regulation of neuron differentiation	3	0.191798	96.29215
Score = 1.03				
GOTERM_BP_FAT	GO:0034754~cellular hormone metabolic process	3	0.063161	63.55895
GOTERM_BP_FAT	GO:0010817~regulation of hormone levels	4	0.080162	72.55087
GOTERM_BP_FAT	GO:0042445~hormone metabolic process	3	0.158085	93.02196
Score = 1.01				
GOTERM_MF_FAT	GO:0046914~transition metal ion binding	27	0.0652	58.48679

Table 3.3 (cont.)

GOTERM_MF_FAT	GO:0046872~metal ion binding	37	0.066955	59.49143
GOTERM_MF_FAT	GO:0043169~cation binding	37	0.074993	63.81308
GOTERM_MF_FAT	GO:0043167~ion binding	37	0.087397	69.65401
GOTERM_MF_FAT	GO:0008270~zinc ion binding	19	0.299808	99.04121

Table 3.4 Genes differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |3|) between low and high activity genotype mice.

Gene*	**Log2 (Control/Selected)**
Foxg1	5.19633
Kcnv1	4.88982
Sp9	4.80959
Dlx1	4.80199
Gpr6	4.71176
Drd1a	4.57751
Zfp831	4.47846
Cpne5	4.47141
Kcnj4	4.42276
Ankrd63	4.39899
Gda	4.30171
Lrrc10b	4.28336
AK040671	4.26887
Kcnf1	4.26674
Chrm1	4.19663
Kcng1	4.01258
Dlx2	3.93618
Rxrg	3.8042
Adora2a	3.75852
Tbr1	3.7204
Rprml	3.60267
Crhbp	3.56026
Rtn4rl2	3.45839
Vdr	3.4123
Actn2	3.34287
Ptprv	3.31905
Kcnh4	3.18327
Neurl1b	3.16774
Sst	3.11108
D430019H16Rik	3.0582
Prss12	3.04173

Table 3.5 Enriched (enrichment score > 3.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes differentially expressed (P-value < 0.001) between low and high activity genotype mice.

Category	Term	Count	PValue	FDR
Score = 5.36				
GOTERM_MF_FAT	GO:0022843~voltage-gated cation channel activity	13	6.43E-09	8.50E-06
GOTERM_MF_FAT	GO:0022832~voltage-gated channel activity	14	3.26E-08	4.31E-05
GOTERM_MF_FAT	GO:0005244~voltage-gated ion channel activity	14	3.26E-08	4.31E-05
GOTERM_MF_FAT	GO:0005261~cation channel activity	15	2.21E-07	2.92E-04
GOTERM_BP_FAT	GO:0030001~metal ion transport	19	3.35E-07	5.21E-04
GOTERM_BP_FAT	GO:0006812~cation transport	20	6.92E-07	0.001077423
GOTERM_MF_FAT	GO:0005249~voltage-gated potassium channel activity	10	7.05E-07	9.32E-04
GOTERM_MF_FAT	GO:0046873~metal ion transmembrane transporter activity	15	1.45E-06	0.00192351
GOTERM_MF_FAT	GO:0022836~gated channel activity	14	5.61E-06	0.007419277
GOTERM_MF_FAT	GO:0005267~potassium channel activity	10	6.46E-06	0.00853665
GOTERM_MF_FAT	GO:0005216~ion channel activity	15	1.24E-05	0.016337178
GOTERM_MF_FAT	GO:0022838~substrate specific channel activity	15	1.75E-05	0.023133726
GOTERM_MF_FAT	GO:0015267~channel activity	15	2.04E-05	0.026973665
GOTERM_MF_FAT	GO:0022803~passive transmembrane transporter activity	15	2.04E-05	0.026973665
GOTERM_BP_FAT	GO:0006811~ion transport	21	2.13E-05	0.033125111
GOTERM_MF_FAT	GO:0030955~potassium ion binding	9	2.64E-05	0.034896362
GOTERM_BP_FAT	GO:0006813~potassium ion transport	10	2.89E-05	0.044941459

Table 3.5 (cont.)

GOTERM_MF_FAT	GO:0031420~alkali metal ion binding	11	4.72E-05	0.062340343
GOTERM_BP_FAT	GO:0015672~monovalent inorganic cation transport	12	2.04E-04	0.316785555
GOTERM_BP_FAT	GO:0055085~transmembrane transport	10	0.038361959	45.59311009

Table 3.6 Genes differentially expressed (FDR-adjusted P-Value < 0.005 and log2(fold change) > |3|) between low and high activity genotype mice.

Gene*	**Log2 (Sedentary/Runner)**
Foxg1	6.58338
Dlx1	6.3096
Kcnv1	5.66069
Drd1a	5.47347
Zfp831	5.30167
Sp9	5.23464
Gpr6	5.23178
Kcnj4	5.10231
Chrm1	5.08345
Ankrd63	5.08332
Gda	5.03023
Krt9	4.95871
Lrrc10b	4.81634
Cpne5	4.77704
Kcnf1	4.77571
Rxrg	4.66121
Kcng1	4.56038
Tbr1	4.18002
Cntnap3	4.17254
Adora2a	4.15213
Crhbp	4.02149
Rtn4rl2	3.9803
Kcnh4	3.69834
Ptprv	3.59163
Lamp5	3.57371
Neurl1b	3.50048
Prss12	3.45094
Rprml	3.39604
Pak6	3.34825
D430019H16Rik	3.33037
Nts	3.26483
Sst	3.24551
Mchr1	3.23157
Actn2	3.15911
Lipg	3.08539
Pde10a	3.05835
Meis2	3.01728

Table 3.7 Enriched (enrichment score > 3.0) clusters of Gene Ontology (GO) biological process (BP), molecular function (MF) Functional Annotation Tool (FAT) categories, and KEGG pathways among the genes differentially expressed (P-value < 0.001) between low and high activity environment mice.

Category	Term	Count	PValue	FDR
Score = 5.43				
Category	Term	Count	PValue	FDR
GOTERM_MF_FAT	GO:0022843~voltage-gated cation channel activity	12	4.97E-09	6.41E-06
GOTERM_MF_FAT	GO:0005249~voltage-gated potassium channel activity	11	5.33E-09	6.88E-06
GOTERM_MF_FAT	GO:0005267~potassium channel activity	12	5.40E-09	6.97E-06
GOTERM_BP_FAT	GO:0006813~potassium ion transport	12	3.03E-08	4.57E-05
GOTERM_MF_FAT	GO:0022832~voltage-gated channel activity	12	1.62E-07	2.09E-04
GOTERM_MF_FAT	GO:0005244~voltage-gated ion channel activity	12	1.62E-07	2.09E-04
GOTERM_MF_FAT	GO:0005261~cation channel activity	13	5.75E-07	7.43E-04
GOTERM_MF_FAT	GO:0022836~gated channel activity	13	2.14E-06	0.002768489
GOTERM_MF_FAT	GO:0046873~metal ion transmembrane transporter activity	13	2.97E-06	0.003839378
GOTERM_MF_FAT	GO:0030955~potassium ion binding	9	4.14E-06	0.005348907
GOTERM_BP_FAT	GO:0015672~monovalent inorganic cation transport	12	1.66E-05	0.025030471

Table 3.7 (cont.)

GOTERM_MF_FAT	GO:0005216~ion channel activity	13	1.94E-05	0.025060351
GOTERM_BP_FAT	GO:0030001~metal ion transport	14	2.57E-05	0.038733152
GOTERM_MF_FAT	GO:0022838~substrate specific channel activity	13	2.64E-05	0.034042037
GOTERM_BP_FAT	GO:0006812~cation transport	15	2.94E-05	0.044264411
GOTERM_MF_FAT	GO:0022803~passive transmembrane transporter activity	13	3.02E-05	0.038976337
GOTERM_MF_FAT	GO:0015267~channel activity	13	3.02E-05	0.038976337
GOTERM_MF_FAT	GO:0031420~alkali metal ion binding	9	2.24E-04	0.289110001
GOTERM_BP_FAT	GO:0006811~ion transport	15	8.49E-04	1.272768078
GOTERM_BP_FAT	GO:0055085~transmembrane transport	7	0.131992523	88.17241872

Figure 3.1 Number of differentially expressed genes that overlap between HH vs LH, HL vs LL, and HH vs HL contrasts.

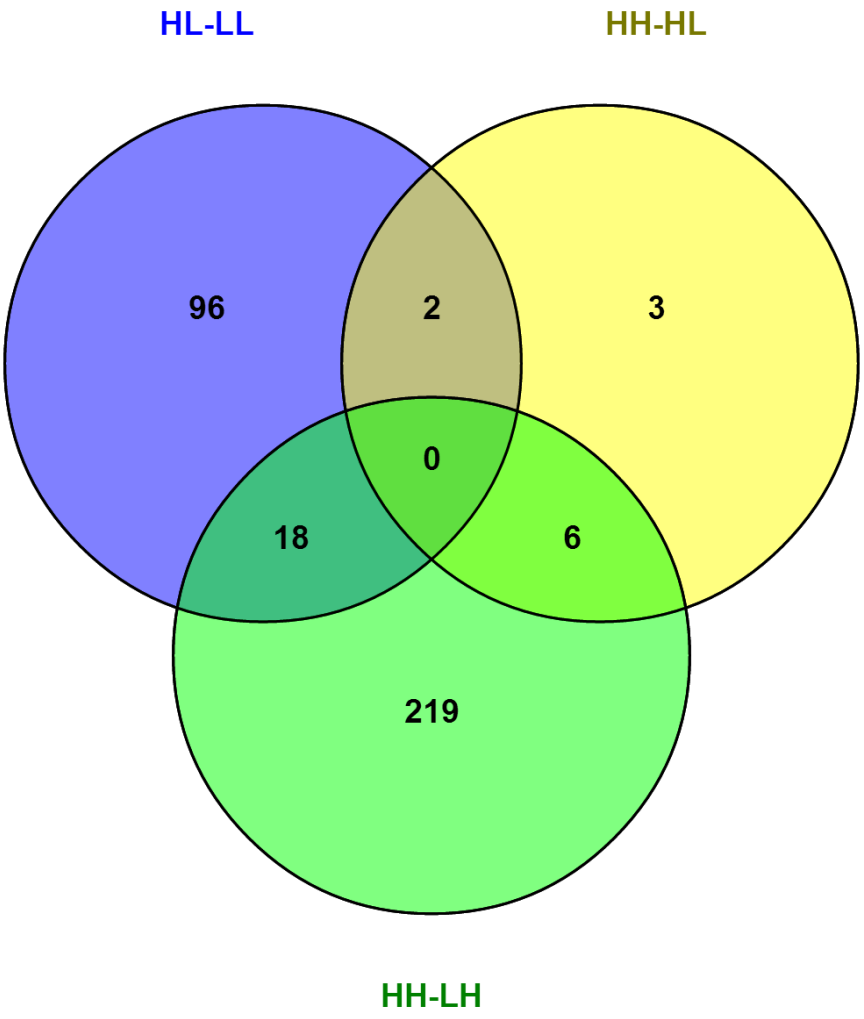


Figure 3.2 Number of differentially expressed genes that overlap between HH vs LH, HL vs LL, and HH vs HL contrasts.

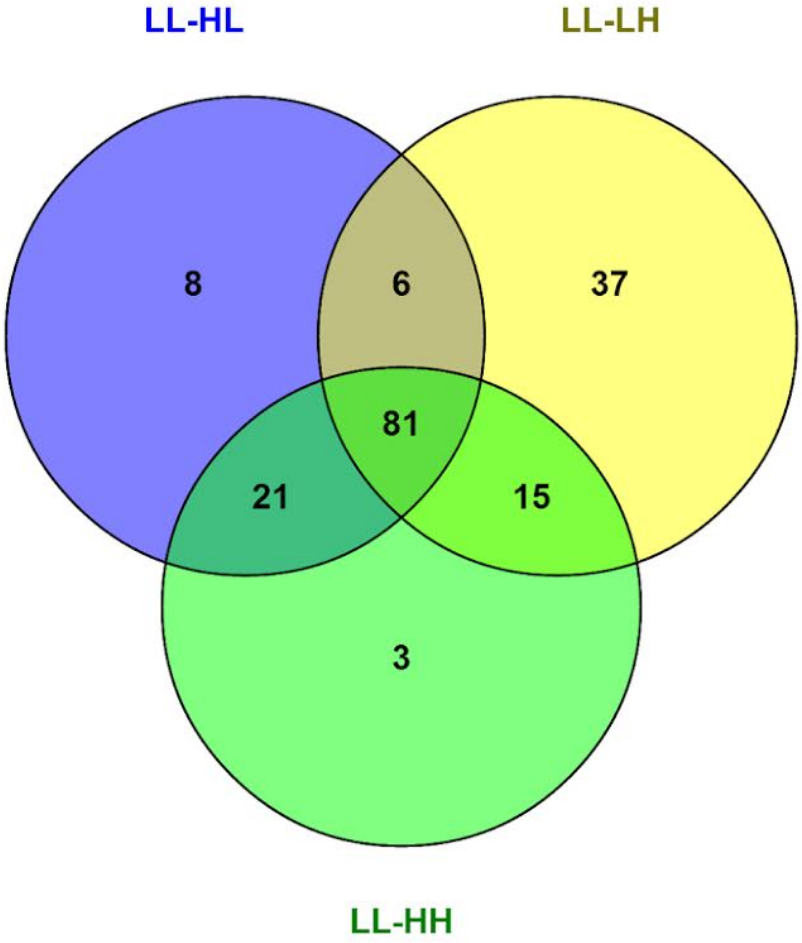


Figure 3.3 Gene network of genes differentially expressed (FDR P-value < 0.05) in orthogonal contrast HL-LL. (Node Color: Red indicates under-expression, Green indicates over-expression, Gray indicates genes connecting differentially expressed gene but not differentially expressed in this study; Node Size: larger indicates a higher fold change, while smaller indicates a smaller fold change

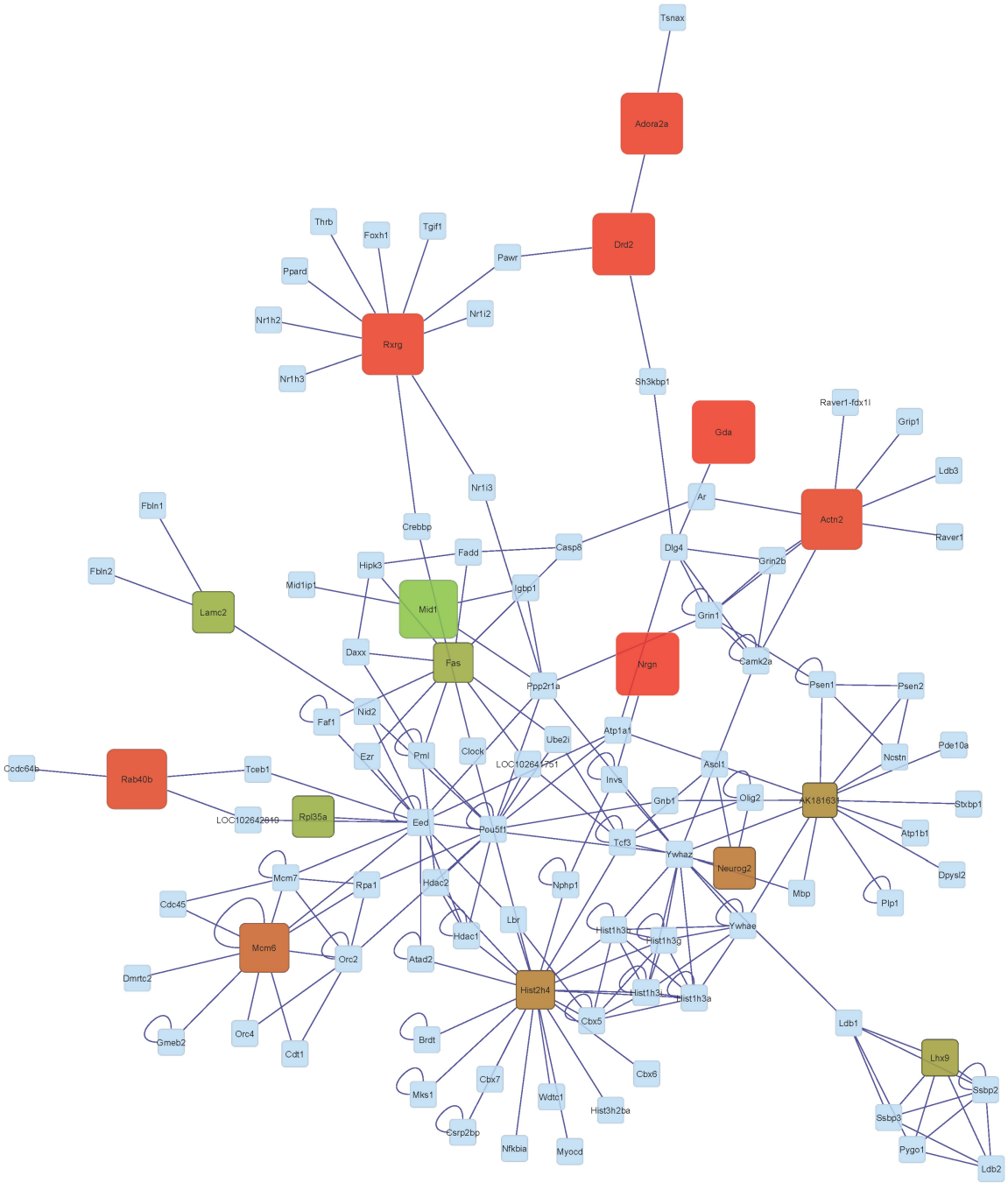


Figure 3.4 Gene network of genes differentially expressed (FDR P-value < 0.05) in orthogonal contrast HH-HL. (Node Color: Red indicates under-expression, Green indicates over-expression, Gray indicates genes connecting differentially expressed gene but not differentially expressed in this study; Node Size: larger indicates a higher fold change, while smaller indicates a smaller fold change

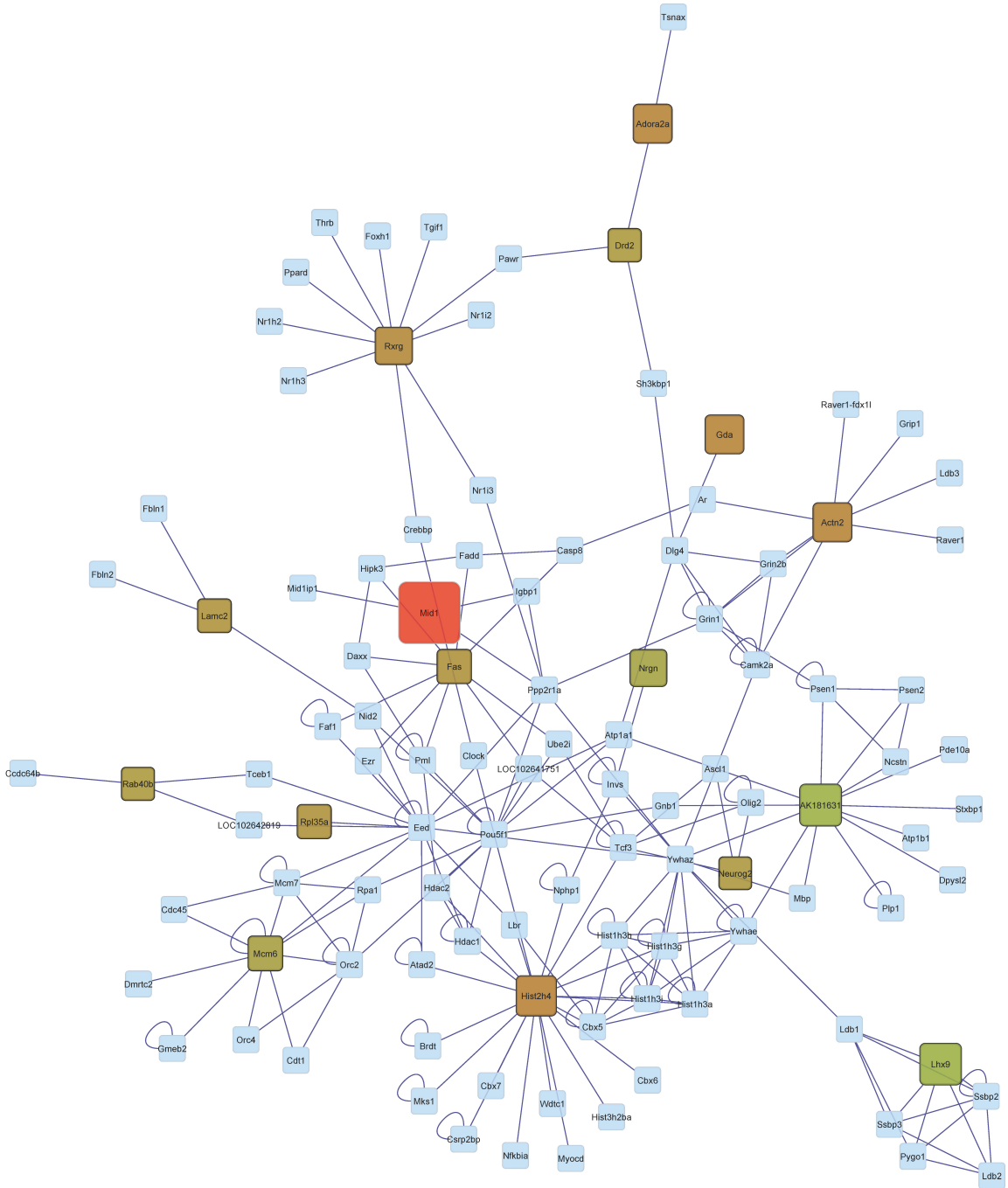
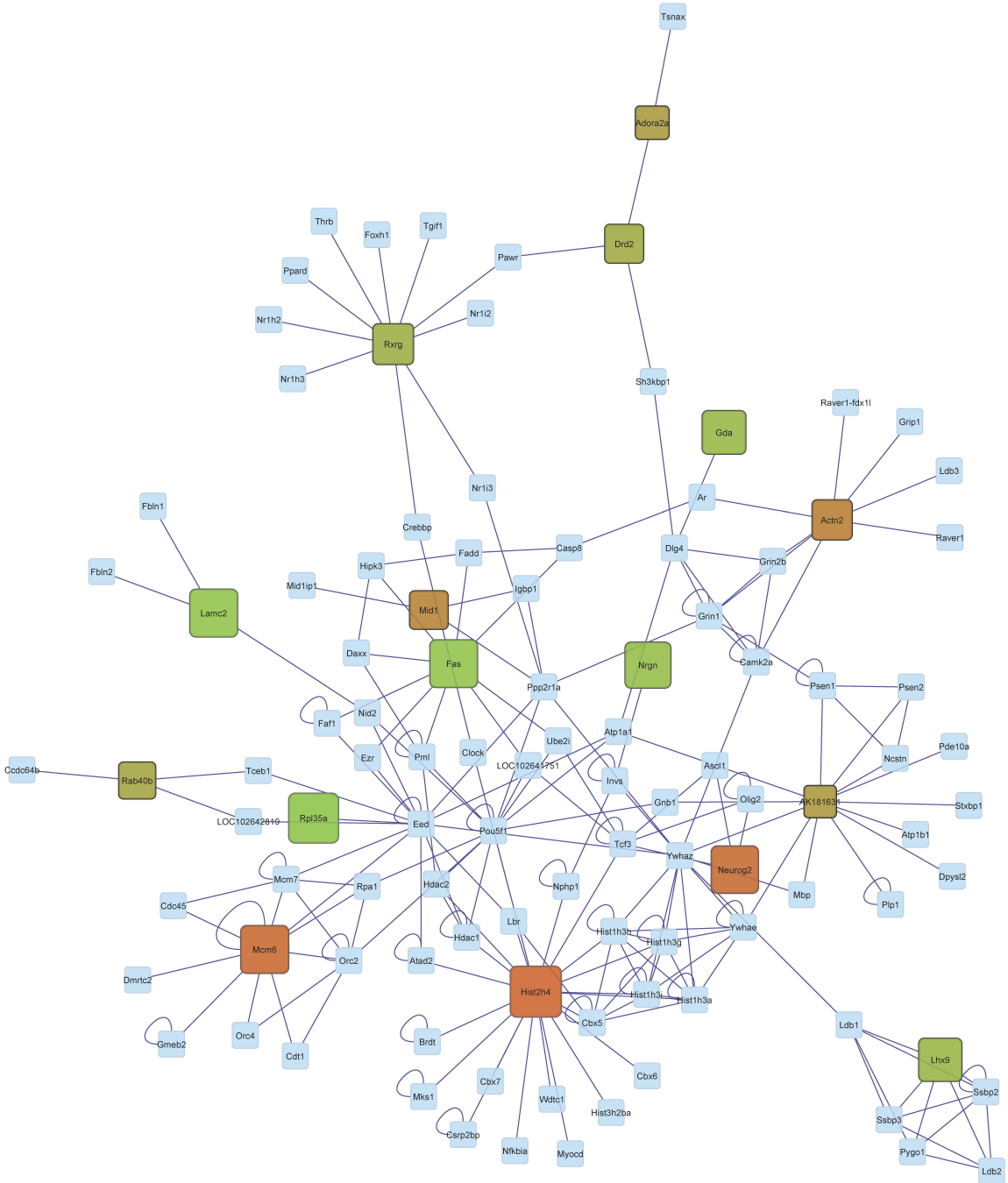


Figure 3.5 Gene network of genes differentially expressed (FDR P-value < 0.05) in orthogonal contrast HH-LH. (Node Color: Red indicates under-expression, Green indicates over-expression, Gray indicates genes connecting differentially expressed gene but not differentially expressed in this study; Node Size: larger indicates a higher fold change, while smaller indicates a smaller fold change



CHAPTER 4. UNCOVERING COMMONALITIES AMONG REWARD-DEPENDENT BEHAVIOR USING GENE NETWORK ANALYSIS

Introduction

Gene pathways have been inferred for addiction to a number of substances including cocaine, amphetamine, morphine, nicotine, alcohol (http://www.genome.jp/kegg/pathway.html). Mouse lines exhibiting high voluntary activity are an established model to study addiction to exercise. A number of genes including dopamine receptors have been associated with high voluntary exercise and other addictive behaviors. Despite ample characterization of genes associated with the high voluntary running model, the similarities between the gene pathways underlying addiction to exercise and to other rewards remains poorly understood.

While standard gene-level and functional analysis allows us to characterize gene expression profiles and enriched functional categories, these analyses do not give us information about interactions between genes. The objective of this study was to uncover commonalities between addiction to high voluntary activity and substances of abuse at the pathway level. Findings from this study could offer insights into gene targets effective in other addictions that could be used to manipulate preference to exercise. Conversely, insights from pathway comparison could aid in the identification of addictions that can be well studied using the high voluntary activity model.

A previous study identified genes differentially expressed between a mouse line selected for high voluntary activity and a control, both on restricted and unrestricted activity environments. From this study, three orthogonal contrasts that provide complementary understanding of the changes in gene expression associated with genotype and environmental differences were studied. These contrasts include: comparison between a mouse line for high voluntary activity and control in restricted and unrestricted activity environments and comparison between restricted and unrestricted activity environment in a high voluntary activity line.

In this chapter, genes differentially expressed in the three contrasts were mapped to the corresponding KEGG pathways of addiction and identified correspondences using

a bipartite graph-theoretic representation of genes and associated functions. The outcome of this bipartite network approach was a novel portrait of the functional makeup of the transcriptome or transcriptomic functionome applied to pathways of addiction.

Materials and Methods

An experiment following a randomized 2 x 2 factorial design including the factors of activity or exercise genetics and environment was carried out. The design included two mouse lines, one line selected for high voluntary physical activity (H) and an unselected control (L) line, and two cage environments enabling two activity levels blocked from wheel running or allowed to run. The two main factors are denoted activity genotype and activity environment respectively. Sample size was n=4/activity genetic-environment combination. Six pairwise contrasts were used to profile the expression patterns involving four groups: high activity genotype-high activity environment (HH); high activity genotype-low activity environment (HL); low activity genotype-high activity environment (LH); and low activity genotype-low activity environment (LL). Among the 6 pairwise contrasts, 3 orthogonal contrasts were of interest: 1) HH vs. LH, 2) HL vs. LL, and 3) HH vs. HL.

Pajek, a software package for analysis and visualization of large networks, was used to for network visualization and analysis on the results of our previous analyses on this data set (Batagelj & Mrvar, 2004). Genes present in at least one of five addiction-related pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa et al., 2008) were considered. Genes in the Cocaine, Amphetamine, Nicotine, Morphine, and Alcoholism KEGG pathways were compiled. This list was combined with the list of genes in each of the three contrasts and associated fold change between combination of activity genotype and environments. These three lists were submitted to Pajek along with information about the associations of the genes to the five addiction-related pathways.

The correspondence between addiction pathways was uncovered using the genes differentially expressed on the three contrasts characterizing exercise addiction using Pajek's bipartite graph-theoretic representation of genes and associated functions. This mapping was decomposed into its dual network projections using mathematical

properties of finite graphs (Chung et al. 1983). This approach was selected because the bipartite graph and its projections describe how function rearrange to unfold modern repertoires of differentially expressed genes. The bipartite network approach offers a depiction of the functional makeup of the transcriptome.

A connection (link or undirected edge) between a function and gene (node) in the bipartite network represents the embedding of that function in the structural scaffold of the corresponding mRNA encoded gene product. Therefore, network connections describe how the elementary alphabet of functions associates with the structural alphabet of transcripts.

Results

Tables 4.1 to 4.3 demonstrate the information used to infer similarities between addiction-related pathways. These tables list the genes characterized in each of the three orthogonal contrasts studied in the high voluntary activity experiment and present in at least four of the five addiction-related pathways studied. Due to space limitations, Tables 4.1 to 4.3 include a limited number of genes that belong to at least four pathways.

Table 4.4 summarizes the number of genes in each addition-related pathway and high voluntary activity orthogonal contrast. The lower number of genes affiliated to the various addiction-related pathways in the HL-LL contrast relative to the other contrast is due to genes that could not be analyzed in that contrasts. Since the HL activity genotype-environment combination was also present in the HH-HL contrast, the fewer genes analyzed in the HL-LL contrast can be traced to genes with very low abundance in the LL activity genotype-environment combination.

Table 4.5 condenses the overlap of genes between pathways and confirms the relationship between contrasts in the numbers of genes associated with each pathway observed in Table 4.4.

Bipartite networks are shown for the contrasts HL-LL, HH-LH, and HH-HL in Figures 4.1, 4.2, and 4.3, respectively. The corresponding waterfall style networks are shown in Figures 4.4, 4.5, and 4.6. The genes in Figures 4.1 to 4.3 are sorted by log2(fold change) from over-expressed (positive log2(fold change)) to under-expressed (negative log2(fold change)). Although the contrasts HH-LH and HH-HL involved the

same genes (Tables 4.4 and 4.5), the bipartite networks demonstrate different gene ordering between these contrasts consistent with the different differential expression observed for the same gene across contrasts. In the waterfall networks shown in Figures 4.4, 4.5, and 4.6, genes are sorted top-down by increasing log2(fold change). Additionally, a parameter was set within Pajek to propel loosely connected genes to the edges and keeps genes shared by multiple pathways organized near the middle.

Discussion

Interpretation of bipartite networks

Tables 4.1 to 4.3 show the number of genes associated with 4 out of 5 of the addiction-related pathways in HL-LL, HH-LH, and HH-HL respectively. Given their high connectivity within the networks, these genes appear to be core addiction-relation genes. Several genes appeared on all three of these lists- Grin3b, Grin2c, Adcy5, Grin2a, Gnas, Grin1, Grin3a, Grin2b, Grin2d, Prkaca. Of these genes, Adenylyl cyclase type 5 (Adcy5; also known as AC5) has been clearly associated with reward processes and addiction to drugs of abuse. Mice lacking the AC5 gene display impaired dopamine D2 receptor function; through increasing activation of the D1 dopamine receptor, administration of haloperidol and clozapine to AC5 knockout mice remarkably increased locomotion, demonstrating this gene's incredible effect on the dopamine receptors and reward (Lee et al., (2002). In AC5 knock out mice, the behavioral effects of morphine including locomotor activation, analgesia, tolerance, reward, and physical dependence/withdrawal symptoms are weakened (Kim et al., 2006). While injection of DAMGO, a mu-opioid receptor selective agonist, into the striatum of normal control mice led to a significant increase in locomotion, so such effect was seen in AC5-knock out mice, which indicates that this reward process-related gene is also involved with activity and locomotion (Kim et al., 2006). CAMP-Dependent Protein Kinase Catalytic Subunit Alpha (Prkaca), a subunit of Adcy5, was found differential expressed as well. GNAS Complex Locus (Gnas) and the several Glutamate receptor (Grin) genes that were found to be significant have not been directly linked to either physical activity or addiction; these genes may serve as potential candidate genes for the study of addiction.

The weak differential expression of the genes listed in Tables 4.1 to 4.3 is not representative of the profile across the complete number of genes studied. These results suggest that genes affiliated to multiple addiction-related pathways do not exhibit major dysregulation on the high voluntary activity pathway.

Table 4.4 displays the number of genes in each contrast group which were found to be associated with multiple pathways. While no genes were found to be associated with all 5 pathways, a small number of genes in each contrast group were associated with 4 out of the 5. These genes are shown in Tables 4.1 to 4.3 for contrasts HL-LL, HH-LH, and HH-HL, respectively. The lower number of genes analyzed in the HL-LL contrast (Table 4.4) was associated to low gene abundance in the LL group. This trend could be linked to The Control genotype in an environment with restricted activity reward could have hinder the expression of some of these genes.

Table 4.5 displays the number of genes in each contrast group associated with each Kegg addiction-related pathway. Interestingly, in each of the 3 contrasts the number of genes associated with the Alcoholism pathway is significantly higher than those of other pathways. This result is intuitive; while years of natural evolution have led to the existence of genes with functions to handle alcohol, while the body has had less time to develop mechanisms to metabolize recently developed synthetic drugs. However, this result may also indicate that addiction to exercise may share molecular mechanisms with alcoholism, and may a suitable model for studies on alcoholism in the future. A noteworthy trend in Table 4.5 is that the number of genes shared across pathways halves with increasing number of pathways. This result highlights the opportunities to identify genes broadly involved in many types of addictions and genes unique to one or few reward-dependent behaviors.

The bipartite networks based on the three contrasts consistently show that the Cocaine and the Alcoholism pathways have genes exhibiting positive, neutral and negative expression patterns in the high voluntary activity experiment. Nicotine had genes with more intermediate expression differential in the two contrasts of activity genotypes (Figures 4.1 and 4.2).

Interpretation of waterfall networks

The waterfall approach to investigate the differentially expressed genes within orthogonal contrast high voluntary activity genotype and environment augmented our understanding of the relationship between activity dependence and four substance-dependence pathways.

Similarities among waterfall networks of substance-dependence pathways between the HL-LL (Figure 4.4) and HH-LH (Figure 4.6) contrasts enabled the uncovering of pathway relationships common to addiction processes independent of activity reward environment. The waterfall networks of pathways have similar horizontal structure in that the pathways in the middle (Amphetamine, Cocaine, and Nicotine) share most genes with the other pathways meanwhile the left- and right-most pathways has a number of genes not shared with the other pathways, yet shared with the high voluntary activity pathway.

Dissimilarities among waterfall networks of substance-dependence pathways between the HL-LL (Figure 4.4) and HH-LH (Figure 4.6) contrasts enabled the uncovering of pathway relationships common to addiction processes dependent of activity reward environment. The waterfall network of pathways corresponding to the genes in HL-LL contrast have fewer connections among pathways than the network in the HH-LH contrast. This is particularly evident in the fewer connections of the Nicotine, Cocaine, and Amphetamine pathways in the HL-LL contrast. This finding indicates that the high voluntary activity pathway in an environment with reward availability (unrestricted running) is more similar to other substance dependence pathways than the same comparison on a reward restricted environment.

Dissimilarities among waterfall networks of substance-dependence pathways between the HL-LL (Figure 4.4) and HH-LH (Figure 4.6) relative to the HH-HL (Figure 4.5) contrast enabled the uncovering of pathway relationships common to addiction processes that could be negated by the activity reward environment. The waterfall networks of pathways have different horizontal structure in the Morphine pathway was less connected to other pathways in the activity genotype contrasts but has relatively more genes in common with the other pathways in the activity environment contrast. This result suggests that the Morphine pathway may be more similar to the pathway of genes differentially expressed in response to activity environment changes.

Similarities among waterfall networks of substance-dependence pathways between the HL-LL (Figure 4.4) and HH-LH (Figure 4.6) relative to the HH-HL (Figure 4.5) contrast enabled the uncovering of pathway relationships common to addiction processes that could be mimicked by the activity reward environment. Alcoholism remains the pathway with fewer genes in common to the pathway of differentially expressed genes associated with activity environment differences, similar to the pathway of differentially expressed genes associated with activity genotype differences.

Significance

This study enabled the visualization of similarities between addiction-related pathways based on gene expression information on the activity addiction pathway. This visualization provide insights into similarities between addiction to drugs of abuse and addiction to activity, as well as provide a foundation for better understanding gene interactions behind the unique phenotype and behavioral traits that selected mice in this experiment display.

With the exception of one study which used bipartite networks created in Pajek to interpret gene expression profiling results and study the human natural killer cell response to Fc receptor activation (Campbell et al., 2015) , the majority of studies using these kinds of networks have investigated social networks rather than genomic information (Jarrett et al., 2015; Lee et al., 2015; Son, Jeong, Kang, Kim, & Lee, 2015). A network approach to gene expression analysis offers the ability to view complex interactions between genes and better visualize often complex results; given the success and clarity of the networks created in this chapter, it is clear that this approach has merit.

TABLES AND FIGURES

Table 4.1 Genes differentially expressed (FDR-adjusted P-value < 0.05) in the HL-LL orthogonal contrast that are associated with 4 out of 5 addiction-related KEGG pathways (Cocaine, Amphetamine, Nicotine, Morphine, and Alcoholism).

Gene name	Log2(fold_change)	P-value	FDR-adjusted P-value	Cocaine	Amphetamine	Nicotine	Morphine	Alcoholism
Grin3b	0.283349	0.520221	0.999983	X	X	X		X
Adcy5	-1.89602	0.0031859	0.18035	X	X		X	X
Grin2a	-0.0406627	0.938633	0.999983	X	X	X		X
Gnas	0.023829	0.950036	0.999983	X	X		X	X
Grin3a	-0.105548	0.847177	0.999983	X	X	X		X
Grin2d	0.0725926	0.903488	0.999983	X	X	X		X
Prkaca	0.0364295	0.928957	0.999983	X	X		X	X
Grin1	0	1	1	X	X	X		X
Grin2b	0	1	1	X	X	X		X

Table 4.2 Genes differentially expressed (FDR-adjusted P-value < 0.05) in the HH-LH orthogonal contrast that are associated with 4 out of 5 addiction-related KEGG pathways (Cocaine, Amphetamine, Nicotine, Morphine, and Alcoholism).

Gene name	:og2(fold_ change)	P_val ue	FDR-adjust ed P-value	Coc aine	Amphet amine	Nico tine	Morp hine	Alcoh olism
Grin3 b	-0.116839	0.598 43	0.999 923	X	X	X		X
Grin2 c	0	1	1	X	X	X		X
Adcy 5	0.175005	0.264 837	0.999 923	X	X		X	X
Grin2 a	0.087299	0.660 92	0.999 923	X	X	X		X
Gnas	0.0920964	0.570 954	0.999 923	X	X		X	X
Grin1	0	1	1	X	X	X		X
Grin3 a	0.0608015	0.767 884	0.999 923	X	X	X		X
Grin2 b	0.468909	0.398 417	0.999 923	X	X	X		X
Grin2 d	0.0409913	0.858 449	0.999 923	X	X	X		X
Prkac a	0.0244129	0.864 523	0.999 923	X	X		X	X

Table 4.3 Genes differentially expressed (FDR-adjusted P-value < 0.05) in the HH-HL orthogonal contrast that are associated with 4 out of 5 addiction-related KEGG pathways (Cocaine, Amphetamine, Nicotine, Morphine, and Alcoholism).

Gene name	:og2(fold_change)	P_value	FDR-adjusted P-value	Cocaine	Amphetamine	Nicotine	Morphine	Alcoholism
Grin3b	-0.0548958	0.783101	0.999891	X	X	X		X
Grin2c	0	1	1	X	X	X		X
Adcy5	-0.106203	0.515129	0.999891	X	X		X	X
Grin2a	0.137666	0.52834	0.999891	X	X	X		X
Gnas	0.0495467	0.766664	0.999891	X	X		X	X
Grin1	0	1	1	X	X	X		X
Grin3a	0.0174095	0.937886	0.999891	X	X	X		X
Grin2b	-0.194388	0.701793	0.999891	X	X	X		X
Grin2d	-0.0874407	0.723722	0.999891	X	X	X		X
Prkaca	0.00369336	0.980037	0.999891	X	X		X	X

Table 4.4 Number of genes significantly expressed in each orthogonal contrast found in KEGG addiction-related pathways (Cocaine, Amphetamine, Nicotine, Morphine, and Alcoholism)

Genes present in:	HL-LL	HH-LH	HH-HL
Cocaine Pathway	40	47	47
Amphetamine Pathway	51	63	63
Nicotine Pathway	36	40	40
Morphine Pathway	73	88	88
Alcoholism Pathway	128	149	149

[1] Four groups: high activity genotype-high activity environment (HH); high activity genotype-low activity environment (HL); low activity genotype-high activity environment (LH); and low activity genotype-low activity environment (LL).

Table 4.5 The number of genes in each orthogonal contrasts shared between multiple KEGG addiction-related pathways (Cocaine, Amphetamine, Nicotine, Morphine, and Alcoholism)

Genes present in:	HL-LL	HH-LH	HH-HL
Only 1 Pathway	125	158	158
2 Pathways	52	57	57
3 Pathways	21	25	25
4 Pathways	9	10	10
5 Pathways	0	0	0

Figure 4.1 Gene-pathway bipartite network constructed with gene expression profiles from the orthogonal contrast HL-LL and addiction-related pathway information from the KEGG database. Nodes are sorted top-down in decreasing order of log2(fold change).

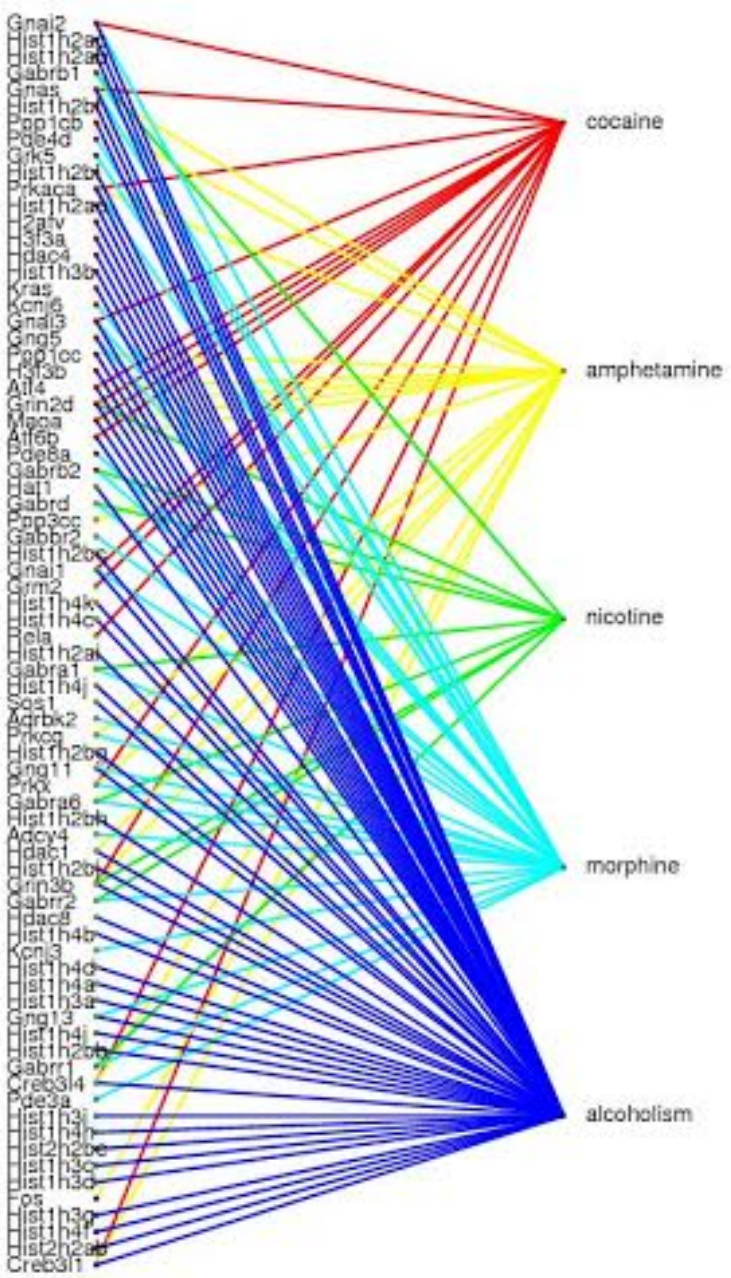


Figure 4.2 Gene-pathway bipartite network constructed with gene expression profiles from orthogonal contrast HH-HL and addiction-related pathway information from the KEGG database. Nodes are sorted top-down in decreasing order of log2(fold change).

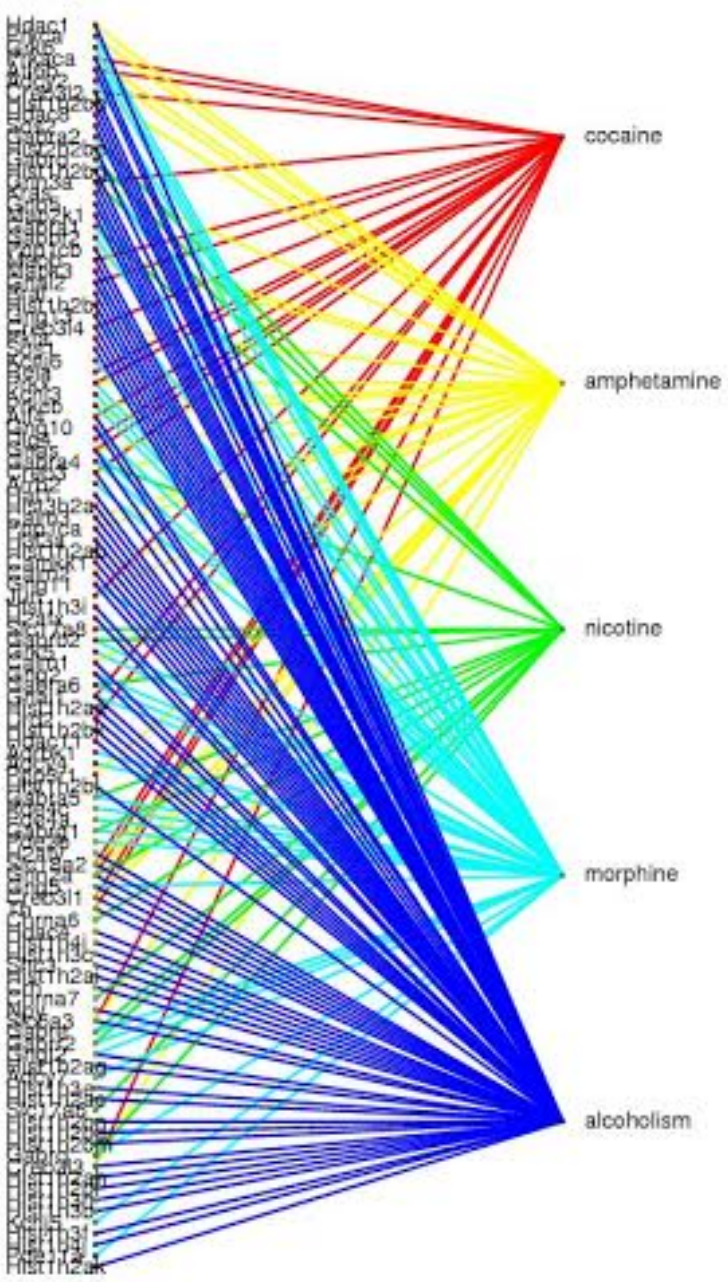


Figure 4.3 Gene-pathway bipartite network constructed with gene expression profiles from orthogonal contrast HH-LH and addiction-related pathway information from the KEGG database. Nodes are sorted top-down in decreasing order of log2(fold change).

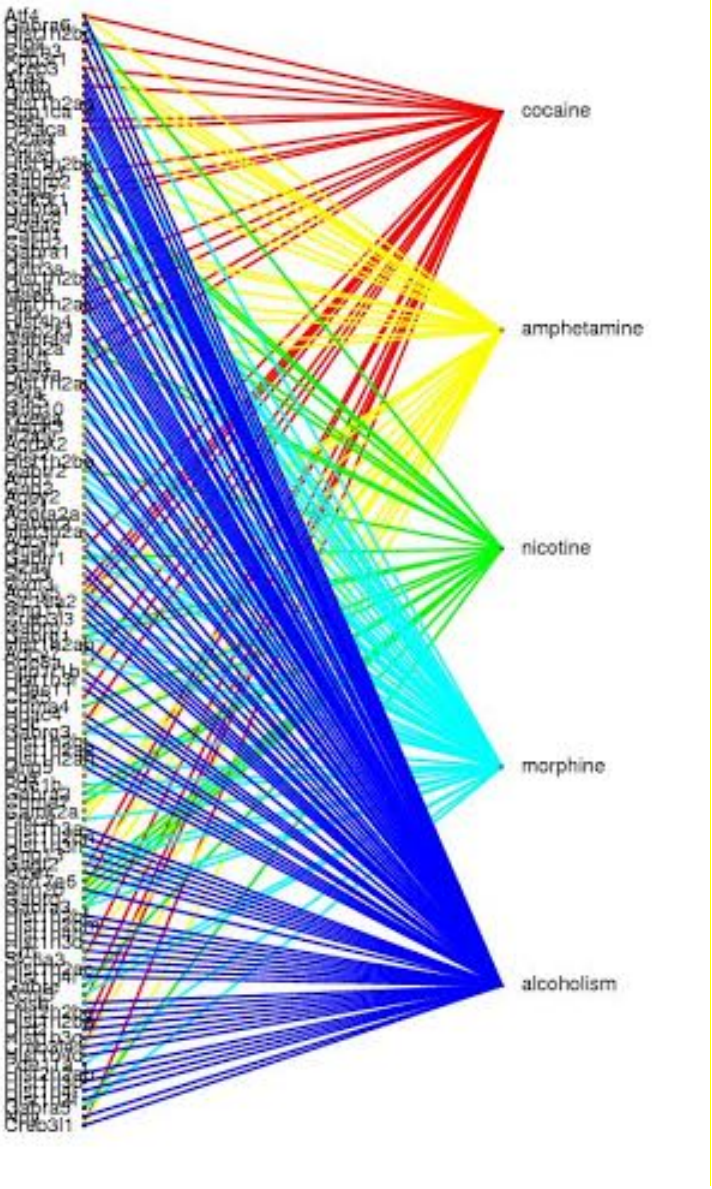


Figure 4.4 Gene-pathway waterfall network constructed with gene expression profiles from orthogonal contrast HL-LL and addiction-related pathway information from the KEGG database.

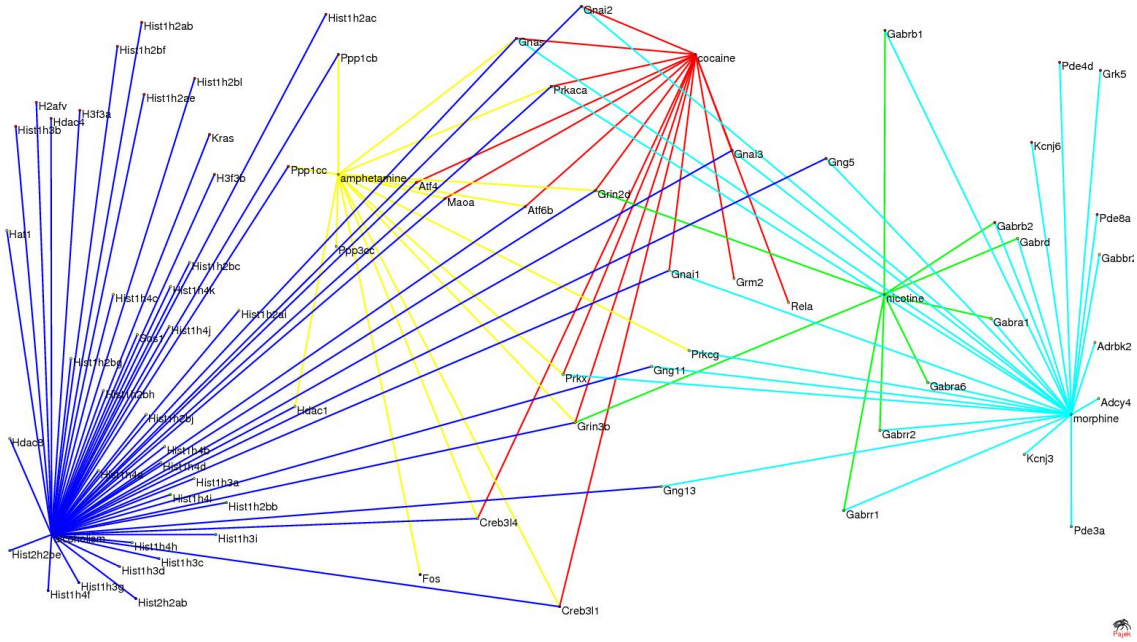


Figure 4.5 Gene-pathway waterfall network constructed with gene expression profiles from orthogonal contrast HH-HL and addiction-related pathway information from the KEGG database

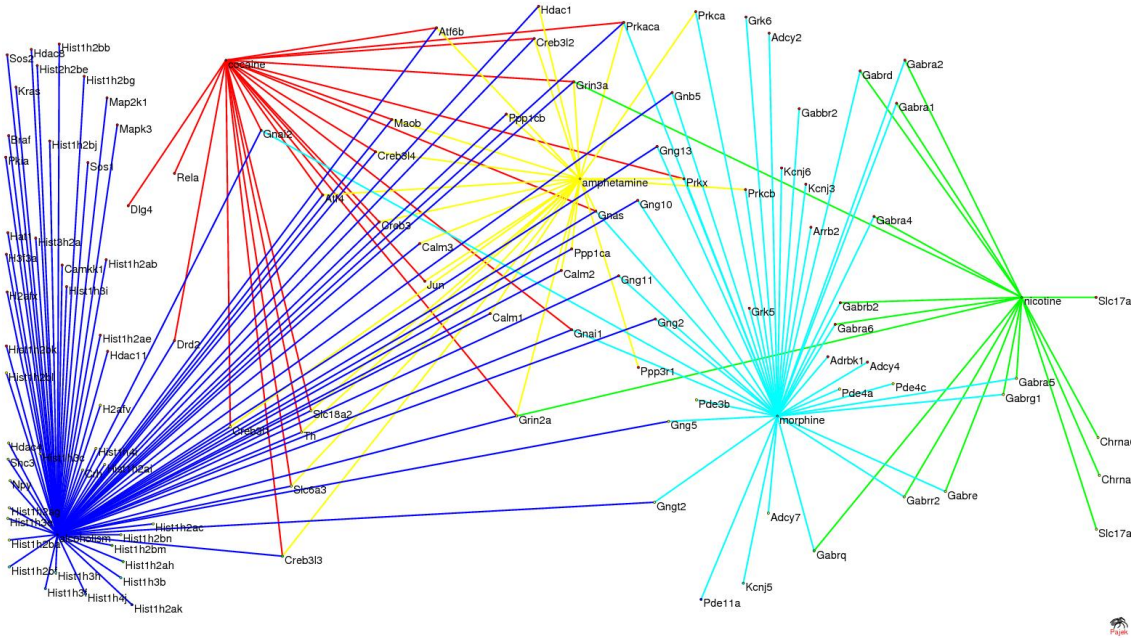
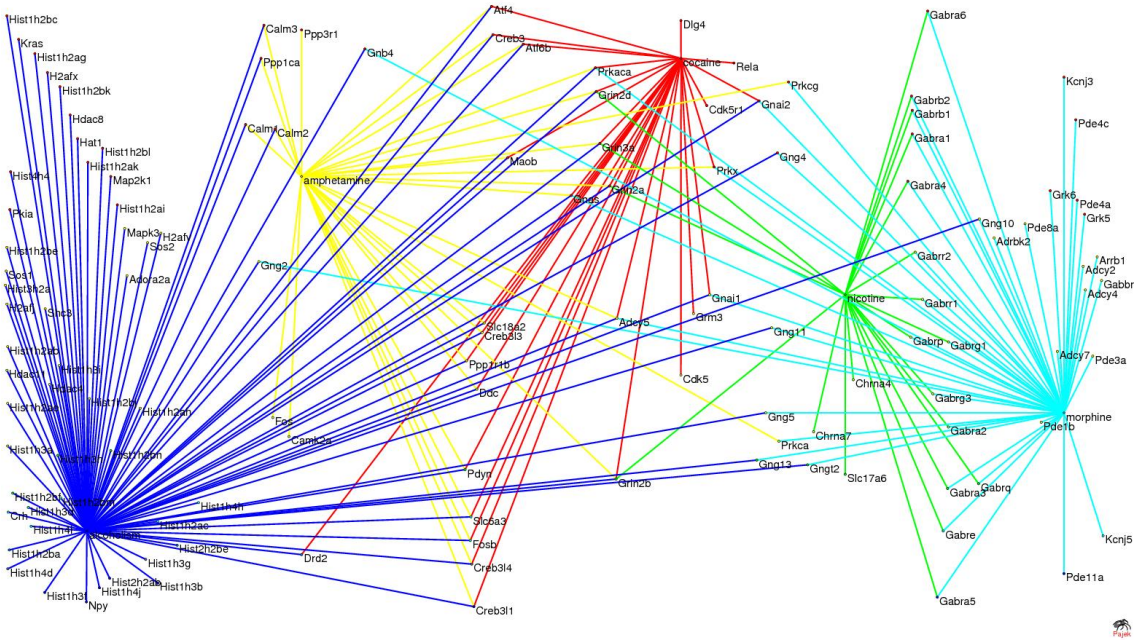


Figure 4.6 Gene-pathway waterfall network constructed with gene expression profiles from orthogonal contrast HH-LH and addiction-related pathway information from the KEGG database



REFERENCES

Adair TH, Gay WJ, Montani JP (1990) Growth regulation of the vascular system: evidence for a metabolic hypothesis. Am J Physiol 259: R393–404.

Alexander GE, DeLong MR & Strick PL. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience, 9, 357-381. doi:10.1146/annurev.ne.09.030186.002041 [doi]

Alliey-Rodriguez N, Zhang D, Badner JA, Lahey BB, Zhang X, Dinwiddie S, . . . The Bipolar Genome Study. (2011). Genome-wide association study of personality traits in bipolar patients. Psychiatric Genetics, 21(4), 190-194. doi:10.1097/YPG.0b013e3283457a31 [doi]

Allison DB, Gadbury GL, Heo M, Ferneandez JR, Lee C-K, et al. (2002) A mixture model approach for the analysis of microarray gene expression data. Comp Stat Data Anal 39: 1–20.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Andersson E, Jensen JB, Parmar M, Guillemot F & Bjorklund A. (2006). Development of the mesencephalic dopaminergic neuron system is compromised in the absence of neurogenin 2. Development (Cambridge, England), 133(3), 507-516. doi:dev.02224 [pii]

Andruchov O, Andruchova O, Wang Y, Galler S (2006) Dependence of cross-bridge kinetics on myosin light chain isoforms in rabbit and rat skeletal muscle fibres. J Physiol 571: 231–242.

Aoyama S, Kase H & Borrelli E. (2000). Rescue of locomotor impairment in dopamine D2 receptor-deficient mice by an adenosine A2A receptor antagonist. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 20(15), 5848-5852. doi:20/15/5848 [pii]

Arber S, Caroni P (1996) Specificity of single LIM motifs in targeting and LIM/LIM interactions in situ. Genes Dev. 10(3):289-300.

Arber S, Halder G, Caroni P (1994) Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. Cell 79: 221–231.

Arber S, Hunter JJ, Ross J Jr, Hongo M, Sansig G, et al. (1997) MLP-deficient mice exhibit a disruption of cardiac cytoarchitectural organization, dilated cardiomyopathy, and heart failure. Cell 88(3):393-403.

Arias-Carrión O, Stamelou M, Murillo-Rodriguez E, Meneendez-González M and Pöpel E. Dopaminergic reward system: a short integrative review. Intl. Arch. Med., 3:24, 2010.

Arnold AP (2009). The organizational-activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. Hormones and Behavior, 55(5), 570-578. doi:10.1016/j.yhbeh.2009.03.011 [doi]

Arnold AP & Breedlove SM (1985). Organizational and activational effects of sex steroids on brain and behavior: A reanalysis. Hormones and Behavior, 19(4), 469-498.

Ashburner M, Ball CA, Blake JA, et al. (2000) Gene Ontology: tool for the unification of biology. Nat Genet 25: 25–29.

Ateia Elmabsout A, Kumawat A, Saenz-Mendez P, Krivospitskaya O, Savenstrand H, et al. (2012) Cloning and Functional Studies of a Splice Variant of CYP26B1 Expressed in Vascular Cells. PLoS ONE 7: e36839.

Balleine BW, Delgado MR & Hikosaka O. (2007). The role of the dorsal striatum in reward and decision-making. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 27(31), 8161-8165. doi:27/31/8161 [pii]

Bal NC, Maurya SK, Sopariwala DH, Sahoo SK, Gupta SC, et al. (2012) Sarcolipin is a newly identified regulator of muscle-based thermogenesis in mammals. Nat Med 18: 1575–1579.

Barash IA, Mathew L, Lahey M, Greaser ML, Lieber RL (2005) Muscle LIM protein plays both structural and functional roles in skeletal muscle. Am J Physiol Cell Physiol 289:C1312–C1320.

Bartek J and Lukas J (2001) Pathways governing G1/S transition and their response to DNA damage. FEBS Lett 490: 117–122.

Batagelj V, Mrvar A. (2004) Pajek - Analysis and visualization of large networks. Math Visual ;95:77–103

Belke TW & Garland T. (2007) A brief opportunity to run does not function as a reinforcer for mice selected for high daily wheel-running rates. Journal of the Experimental Analysis of Behavior, 88(2), 199-213. doi:10.1901/jeab.2007.62-06

Beninger RJ, Hoffman DC & Mazurski EJ (1989). Receptor subtype-specific dopaminergic agents and conditioned behavior. Neuroscience and Biobehavioral Reviews, 13(2-3), 113-122. doi:S0149-7634(89)80019-3 [pii]

Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 289-300, 1995.

Bera TK, Liu X-F, Yamada M, Gavrilova O, Mezey E, Tessarollo L, Anver M, Hahn Y, Lee B, and Pastan I. A model for obesity and gigantism due to disruption of the Ankrd26 gene. Proceedings of the National Academy of Sciences 105: 270-275, 2008.

Berchtold MW, Brinkmeier H & Muntener M. (2000) Calcium ion in skeletal muscle: Its crucial role for muscle function, plasticity, and disease. Physiological Reviews, 80(3), 1215-1265.

Bernardo BL, Wachtmann TS, Cosgrove PG, Kuhn M, Opsahl AC, et al. (2010) Postnatal PPARdelta activation and myostatin inhibition exert distinct yet complementary effects on the metabolic profile of obese insulin-resistant mice. PLoS ONE 5: e11307.

Berridge KC and Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Brain Res. Rev., 28(3):309-369, 1998.

Birkhead TR, Hatchwell BJ, Lindner R, Blomqvist D, Pellatt EJ, Griffiths R, and Lifjeld JT. Extra-pair paternity in the Common Murre. The Condor 103: 158-162, 2001.

Blackburn J and Mansell JP. The emerging role of lysophosphatidic acid (LPA) in skeletal biology. Bone 50: 756-762, 2012.

Blaise S, De Parseval N, and Heidmann T. Functional characterization of two newly identified Human Endogenous Retrovirus coding envelope genes. Retrovirology 2: 19, 2005.

Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, et al. (2010) Manipulation of FASTQ data with Galaxy. Bioinformatics 26: 1783–1785.

Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, et al. (2011) Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol Chapter 19 Unit 19. 10: 11–21.

Bowman EM, Aigner TG & Richmond BJ (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. Journal of Neurophysiology, 75(3), 1061-1073.

Brown RM & Short JL (2008). Adenosine A(2A) receptors and their role in drug addiction. The Journal of Pharmacy and Pharmacology, 60(11), 1409-1430. doi:10.1211/jpp/60.11.0001 [doi]

Brown RR, Ozaki Y, Datta SP, Borden EC, Sondel PM & Malone DG. (1991) Implications of interferon-induced tryptophan catabolism in cancer, auto-immune diseases and AIDS. Advances in Experimental Medicine and Biology, 294, 425-435.

Brown RR, Ozaki Y, Datta SP, Borden EC, Sondel PM and Malone DG. (1991) Implications of interferon-induced tryptophan catabolism in cancer, auto-immune diseases and AIDS. Adv. Exp. Med. Biol., 294:425-435.

Bullard JH, Purdom E, Hansen KD and Dudoit S. (2010) Evaluation of statistical methods for normalization and differential expression in mRNA-seq experiments. BMC Bioinformatics, 11:94.

Caetano-Anollés K, Mishra S & Rodriguez-Zas, SL (2015). Synergistic and antagonistic interplay between myostatin gene expression and physical activity levels on gene expression patterns in triceps brachii muscles of C57/BL6 mice - Public Library of Science. doi:10.1371/journal.pone.0116828

Campbell AR, Regan K, Bhave N, Pattanayak A, Parihar R, Stiff AR, Carson WE. (2015). Gene expression profiling of the human natural killer cell response to fc receptor

activation: Unique enhancement in the presence of interleukin-12. BMC Medical Genomics, 8(1), 66-015-0142-9. doi:10.1186/s12920-015-0142-9 [doi]

Carelli RM, King VC, Hampson RE & Deadwyler SA. (1993) Firing patterns of nucleus accumbens neurons during cocaine self-administration in rats. Brain Research, 626(1-2), 14-22. doi:0006-8993(93)90557-4.

Cassar-Malek I, Passelaigue F, Bernard C, Léger J, Hocquette JF (2007) Target genes of myostatin loss-of-function in muscles of late bovine fetuses. BMC Genomics 8, 63.

Chang JY, Sawyer SF, Lee RS and Woodward DJ (1994). Electrophysiological and pharmacological evidence for the role of the nucleus accumbens in cocaine self-administration in freely moving rats. J. Neurosci. 14, 1224–1244.

Chen Y, Lun AT, and Smyth GK. Differential Expression Analysis of Complex RNA-seq Experiments Using edgeR. 2014.

Choi JW, Herr DR, Noguchi K, Yung YC, Lee C-W, Mutoh T, Lin M-E, Teo ST, Park KE, and Mosley AN. LPA receptors: subtypes and biological actions. Annual review of pharmacology and toxicology 50: 157-186, 2010.

Choudhary D, Jansson I, Stoilov I, Sarfarazi M, Schenkman JB (2005) Expression patterns of mouse and human CYP orthologs (families 1–4) during development and in different adult tissues. Arch Biochem Biophys 436: 50–61.

Chung FRK, Erdös P, Spencer J (1983) On the decomposition of graphs into complete bipartite subgraphs. In: Erdös P, Alpar L, Halasz G, Saeközy A (eds) Studies in pure mathematics, To the memory of Paul Turán. Bikhäuser Verlag, Budapest, Hungary, pp. 95-101

Comings DE, Rosenthal RJ, Lesieur HR, Rugle LJ, Muhleman D, et al. 1996. A study of the dopamine D2 receptor gene in pathological gambling. Pharmacogenetics 6:223–34

Craciunescu CN, Albright CD, Mar M-H, Song J, and Zeisel SH. Choline availability during embryonic development alters progenitor cell mitosis in developing mouse hippocampus. The Journal of nutrition 133: 3614-3618, 2003.

Cui HX, Liu RR, Zhao GP, Zheng MQ, Chen JL, Wen J. (2012) Identification of differentially expressed genes and pathways for intramuscular fat deposition in

pectoralis major tissues of fast-and slow-growing chickens. BMC genomics 13: 213.

Cummins B & Salmons S (1999) Changes in the synthesis of total proteins induced by chronic electrical stimulation of skeletal muscle. Basic Appl Myol 9: 19–28.

Davis MJ, Donovitz JA, Hood JD (1992) Stretch-activated single-channel and whole cell currents in vascular smooth muscle cells. Am J Physiol Cell Physiol 262: C1083–C1088.

Da Wei Huang BTS & Lempicki RA. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols 4: 44-57.

Dawson DA, Darby S, Hunter FM, Krupa AP, Jones IL, and Burke T. (2001) A critique of avian CHD‐based molecular sexing protocols illustrated by a Z‐chromosome polymorphism detected in auklets. Molecular Ecology Notes 1: 201-204.

de Boer J, Andressoo JO, de Wit J, Huijmans J, Beems RB, et al. (2002) Premature aging in mice deficient in DNA repair and transcription. Science 296: 1276–1279.

De Bruin LA, Schasfoort EM, Steffens AB & Korf J. (1990). Effects of stress and exercise on rat hippocampus and striatum extracellular lactate. The American Journal of Physiology, 259(4 Pt 2), R773-9.

Del Fabbro C, Scalabrin S, Morgante M, Giorgi FM. An extensive evaluation of read trimming effects on Illumina NGS data analysis. PLoS One. 2013;8(12):e85024.

Delfino KR & Rodriguez-Zas SL (2013). Transcription factor-microRNA-target gene networks associated with ovarian cancer survival and recurrence. PloS One, 8(3), e58608. doi:10.1371/journal.pone.0058608 [doi]

Delfino KR, Southey B, Sweedler J, and Rodriguez-Zas S. Genome-wide census and expression profiling of chicken neuropeptide and prohormone convertase genes. Neuropeptides 44: 31-44, 2010.

De Montis M, Devoto P, Meloni D, Saba P & Tagliamonte A. (1992) Decreased adenosine A2A receptor function in morphine dependent rats. Pharmacol. Res., 25, 232–233.

Dennis G, Sherman B, Hosack D, Yang J, Gao W, Lane H, Lempicki R (2003) DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol 4: R60.

de Souza Silva MA, Mattern C, Topic B, Buddenberg TE & Huston JP. (2009) Dopaminergic and serotonergic activity in neostriatum and nucleus accumbens enhanced by intranasal administration of testosterone. European Neuropsychopharmacology : The Journal of the European College of Neuropsychopharmacology, 19(1), 53-63. doi:10.1016/j.euroneuro.2008.08.003 [doi]

Di Chiara G, Tanda G, Bassareo V, Pontieri F, Acquas E, Fenu S, Cadoni C and Carboni E. Drug addiction as a disorder of associative learning. Role of nucleus accumbens shell/extended amygdala dopamine. Ann. N. Y. Acad. Sci., 877:461-485, 1999.

Dlugosz EM, Schutz H, Meek TH, Acosta W, Downs CJ, Platzer EG . . . Garland TJ. (2013) Immune response to a trichinella spiralis infection in house mice from lines selectively bred for high voluntary wheel running. The Journal of Experimental Biology, 216(Pt 22), 4212-4221. doi:10.1242/jeb.087361 [doi]

Dodd ML, Klos KJ, Bower JH, Geda YE, Josephs KA, Ahlskog JE. Pathological gambling caused by drugs used to treat Parkinson disease. Arch Neurol 2005;62(9):1377–81.

Doya K (2000) Complementary roles of basal ganglia and cerebellum in learning and motor control.Curr Opin Neurobiol 10:732–739.

Dubiec A & Zagalska-Neubauer M. Molecular techniques for sex identification in birds. Biological Letters 43: 3-12, 2006.

Duhl DM, Vrieling H, Miller KA, Wolff GL and Barsh GS. (1994)Neomorphic agouti mutations in obese yellow mice. Nature Genet., 8:59-65.

Vivier E, Tomasello E , Baratin M, Walzer T and Ugolini S. (2008) Functions of natural killer cells. Nature Immunol. 9:503-510.

Ecarnot-Laubriet A, De Luca K, Vandroux D, Moisant M, Bernard C, et al. (2000) Downregulation and nuclear relocation of MLP during the progression of right

ventricular hypertrophy induced by chronic pressure overload. J Mol Cell Cardiol 32(12):2385-95.

Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30: 207–210.

Ehler E, Horowits R, Zuppinger C, Price RL, Perriard E, et al. (2001) Alterations at the intercalated disk associated with the absence of muscle LIM protein. J Cell Biol 153(4):763-72.

Ehrenkranz J, Bliss E & Sheard MH (1974). Plasma testosterone: Correlation with aggressive behavior and social dominance in man. Psychosomatic Medicine, 36(6), 469-475.

Ellegren H. (1996) First gene on the avian W chromosome (CHD) provides a tag for universal sexing of non-ratite birds. Proceedings of the Royal Society of London Series B: Biological Sciences 263: 1635-1641.

Ellegren H. (2001) Hens, cocks and avian sex determination. EMBO reports 2: 192-196.

Elmabsout AA, Kumawat A, Saenz-Mendez P, Krivospitskaya O, Savenstrand H, Olofsson PS, Eriksson LA, Strid A, Valen G, Torma H.and Sirsjö A. (2012) Cloning and Functional Studies of a Splice Variant of CYP26B1 Expressed in Vascular Cells. PLoS ONE, 7:e36839.

Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE, The Mouse Genome Database Group (2014) The Mouse Genome Database (MGD): facilitating mouse as a model of human biology and disease. Nucleic Acids Res. 43:D726-D736.

Fajardo VA, Bombardier E, Vigna C, Devji T, Bloemberg D, et al. (2013) Co-expression of SERCA isoforms, phospholamban and sarcolipin in human skeletal muscle fibers. PLoS ONE 8: e84304.

Flick MJ, Konieczny SF (2000) The muscle regulatory and structural protein MLP is a cytoskeletal binding partner of betaI-spectrin. J Cell Sci 113 (Pt 9)():1553-64.

Franklin KB & Vaccarino FJ. (1983) Differential effects of amphetamine isomers on SN self-stimulation: Evidence for DA neuron subtypes. Pharmacology, Biochemistry, and Behavior, 18(5), 747-751.

Garcia-Fuster MJ, Ferrer-Alcan M, Miralles A & Garcia-Sevilla J. (2003) Modulation of fas receptor proteins and dynamin during opiate addiction and induction of opiate withdrawal in rat brain. Naunyn-Schmiedeberg's Archives of Pharmacology, 368(5), 421-431. doi:10.1007/s00210-003-0801-9

Garland T & Freeman PW (2005). Selective breeding for high endurance running increases hindlimb symmetry. Evolution, 59(8), 1851-1854. doi:10.1111/j.0014-3820.2005.tb01832.x

Garland T, Schutz H, Chappell MA, Keeney BK, Meek TH, Copes LE . . . Eisenmann, JC (2010). The biological control of voluntary exercise, spontaneous physical activity and daily energy expenditure in relation to obesity: Human and rodent perspectives. Journal of Experimental Biology, 214(2), 206-229.

Gat-Viks I, Sharan R and Shamir R. Scoring clustering solutions by their biological relevance. Bioinformatics 19: 2381-2389, 2003.

Gerber DJ Sotnikova TD, Gainetdinov RR, Huang SY, Caron MG & Tonegawa S. (2001) Hyperactivity, elevated dopaminergic transmission, and response to amphetamine in M1 muscarinic acetylcholine receptor-deficient mice. Proceedings of the National Academy of Sciences of the United States of America, 98(26), 15312-15317. doi:261583798 [pii]

Giannesini B, Vilmen C, Amthor H, Bernard M, Bendahan D (2013) Lack of myostatin impairs mechanical performance and ATP cost of contraction in exercising mouse gastrocnemius muscle in vivo. Am J Physiol Endocrinol Metab 305: E33–40.

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, et al. (2005) Galaxy: a platform for interactive large-scale genome analysis. Genome Res 15: 1451–1455.

Gineste C, De Winter JM, Kohl C, Witt CC, Giannesini B, et al. (2013) In vivo and in vitro investigations of heterozygous nebulin knock-out mice disclose a mild skeletal muscle phenotype. Neuromuscul Disord 23: 357–369.

Goecks J, Nekrutenko A, Taylor J and The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol., 11(8):R86, 2010.

Goldfarb AH, Hatfield BD, Sforzo GA, Flynn MG (1987) Serum beta-endorphin levels during a graded exercise test to exhaustion. Med Sci Sports Exerc 19:78-82.

Gozes I, Werner H, Fawzi M, Abdelatty A, Shani Y, Fridkin M, and Koch Y. Estrogen regulation of vasoactive intestinal peptide mRNA in rat hypothalamus. Journal of Molecular Neuroscience 1: 55-61, 1989.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, and Zeng Q. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nature biotechnology 29: 644-652, 2011.

Griffiths R, Daan S, and Dijkstra C. Sex identification in birds using two CHD genes. Proceedings of the Royal Society of London Series B: Biological Sciences 263: 1251-1256, 1996.

Grobet L, Martin LJ, Poncelet D, Pirottin D, Brouwers B, et al. (1997) A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. Nat Genet 17: 71–74.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8: 1494–1512. doi: 10.1038/nprot.2013.084

Hamrick MW, McPherron AC, Lovejoy CO (2002) Bone mineral content and density in the humerus of adult myostatin-deficient mice. Calcif Tissue Int 71: 63–68.

Hannon RM,Meek TH, Acosta W, Maciel RC, Schutz H & Garland,T. (2010) Sex-specific heterosis in line crosses of mice selectively bred for high locomotor activity. Behavior Genetics, 41(4), 615-624. doi:10.1007/s10519-010-9432-3

Harber VJ, Sutton JR. (1984) Endorphins and exercise. Sports Medicine 1: 154171.

Harper BD, Beckerle MC, Pomiès P (2000) Fine mapping of the alpha-actinin binding site within cysteine-rich protein. Biochem J. 350 Pt 1():269-74.

Hausenblas HA & Downs DS. (2002) Exercise dependence: a systematic review. Psychol. Sport Exerc., 3:89-123.

Hayamizu TF, Wicks MN, Davidson DR, Burger A, Ryngwald M, Baldock RA (2013) J. Biomed. Semantics 4:15.

Hayashi S, McMahon A. (2002) Efficient Recombination in Diverse Tissues by a Tamoxifen-Inducible Form of Cre: A Tool for Temporally Regulated Gene

Activation/Inactivation in the Mouse. Developmental Biology 244 (2): 305–318. doi:10.1006/dbio.2002.0597. ISSN 0012-1606.

Heineke J, Auger-Messier M, Xu J, Sargent M, York A, Welle S, et al. (2010) Genetic deletion of myostatin from the heart prevents skeletal muscle atrophy in heart failure. Circulation 121: 419–425.

Heng YW & Koh CG. (2010) Actin cytoskeleton dynamics and the cell division cycle. Int J Biochem Cell Biol 42: 1622–1633.

Hinoki A, Kimura K, Higuchi S, Eguchi K, Takaguri A, et al. (2010) p21-activated kinase 1 participates in vascular remodeling in vitro and in vivo. Hypertension 55: 161–5.

Hinshaw W, Burmester B, Creamer A, Hess C, Howes J, Insko W, and Wilson W. Coturnix (Coturnix coturnix japonica): standards and guidelines for the breeding, care and management of laboratory animals. National Academy of Sciences, Washington DC 1969.

Hollerman JR & Schultz W. (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat. Neurosci. 1, 304–309.

Homma K, Siopes TD, Wilson WO, and McFarland LZ. Identification of sex of day-old quail (Coturnix coturnix japonica) by cloacal examination. Poultry Science 45: 469-472, 1966.

Hommer DW, Bjork JM and Gillman JM. (2011) Imaging brain response to reward in addictive disorders. Ann N Y Acad Sci. 1216:50–61

Hoshi E, Tremblay L, Féger J, Carras PL, Strick PL (2005) The cerebellum communicates with the basal ganglia. Nat Neurosci 8:1491–1493.

Houk JC. (2005) Agents of the mind. Biol Cybern 92:427–437.

Huang da W, Sherman BT & Lempicki RA. (2009a). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research, 37(1), 1-13. doi:10.1093/nar/gkn923

Huang da W, Sherman BT & Lempicki RA. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4(1), 44-57. doi:10.1038/nprot.2008.211

Huang da W, Sherman BT, Tan Q, Collins JR, Alvord WG, et al. (2007) The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 8: R183.

Hughes PE, Alexi T, Williams CE, Clark RG, Gluckman PD (1999) Administration of recombinant human Activin-A has powerful neurotrophic effects on select striatal phenotypes in the quinolinic acid lesion model of Huntington's disease. Neuroscience 92: 197–209.

Hunter DJ. (2005) Gene-environment interactions in human disease. Nature Rev. Genet., 6:287-298.

Hyman SE. (2005) Addiction: A disease of learning and memory. The American Journal of Psychiatry, 162(8), 1414-1422. doi:162/8/1414

Hyman SE. (2005) Addiction: a disease of learning and memory. Am J Psychiatry 162:1414-1422.

Ivlieva N. (2011) Involvement of the mesocorticolimbic dopaminergic system in adaptive behavior. Neurosci Behav Physiol, 41(7), 715-29. doi:10.1007/s11055-011-9477-7

Jarrett N, Davis K, Addington-Hall C, Duke J, Corner S & Lathlean J. (2015). The networks of care surrounding cancer palliative care patients. BMJ Supportive & Palliative Care, doi:bmjspcare-2014-000782 [pii]

Kadi F (2000) Adaptation of human skeletal muscle to training and anabolic steroids. Acta Physiol Scand, Suppl 646: 1–51.

Kahn NW, St. John J, and Quinn TW. Chromosome-specific intron size differences in the avian CHD gene provide an efficient method for sex identification in birds. The Auk 1074-1078, 1998.

Kalivas PW & Volkow ND. (2005) The neural basis of addiction: A pathology of motivation and choice. The American Journal of Psychiatry, 162(8), 1403-1413. doi:162/8/1403

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, et al. (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36: D480–484.

Keeney BK, Meek TH, Middleton KM, Holness LF & Garland TJ (2012) Sex differences in cannabinoid receptor-1 (CB1) pharmacology in mice selectively bred for high voluntary wheel-running behavior. Pharmacology, Biochemistry, and Behavior, 101(4), 528-537. doi:10.1016/j.pbb.2012.02.017 [doi]

Kele J, Simplicio N, Ferri AL, Mira H, Guillemot F, Arenas E & Ang SL. (2006) Neurogenin 2 is required for the development of ventral midbrain dopaminergic neurons. Development (Cambridge, England), 133(3), 495-505. doi:133/3/495 [pii]

Kelly SA Rezende EL, Chappell MA, Gomes FR, Kolb EM, Malisch JL . . . Garland TJ. (2014) Exercise training effects on hypoxic and hypercapnic ventilatory responses in mice selected for increased voluntary wheel running. Experimental Physiology, 99(2), 403-413. doi:10.1113/expphysiol.2013.076018

Killcoyne S, Carter GW, Smith J & Boyle J (2009). Cytoscape: A community-based framework for network modeling. Methods in Molecular Biology (Clifton, N.J.), 563, 219-239. doi:10.1007/978-1-60761-175-2_12 [doi]

Kim KS, Lee KW, Lee KW, Im JW, Yoo JY, Kim SW, Lee JK, Nestler EJ, Han PL. (2006) Adenylyl cyclase type 5 (AC5) is an essential mediator of morphine action. Proc. Natl. Acad. Sci. U.S.A.

Kim YP, Kim H, Shin MS, Chang HK, Jang MH, Shin MC, . . . Kim CJ. (2004) Age-dependence of the effect of treadmill exercise on cell proliferation in the dentate gyrus of rats. Neuroscience Letters, 355(1-2), 152-154. doi:S0304394003013089 [pii]

Kingwell B (2000) Nitric oxide as a metabolic regulator during exercise: effects of training in health and disease. Clin Exp Pharmacol Physiol 27: 239–250.

Kirber MT, Walsh JV & Singer JJ. (1988) Stretch-activated ion channels in smooth muscle: A mechanism for the initiation of stretch-induced contraction. Pflugers Archiv : European Journal of Physiology, 412(4), 339-345.

Kirber MT, Walsh JV, Singer JJ (1988) Stretch-activated ion channels in smooth muscle: a mechanism for the initiation of stretch-induced contraction. Pflügers Arch 412: 339–345.

Klos KJ, Bower JH, Josephs KA, Matsumoto JY, Ahlskog JE. Pathological hypersexuality predominantly linked to adjuvant dopamine agonist therapy in

Parkinson's disease and multiple system atrophy. Parkinsonism Relat Disord 2005;11(6):381–6.

Knoll R, Hoshijima M, Chien KR. (2002) Muscle LIM protein in heart failure. Exp Clin Cardiol 7(2-3):104-5.

Kolb EM, Kelly SA & Garland TJ. (2013) Mice from lines selectively bred for high voluntary wheel running exhibit lower blood pressure during withdrawal from wheel access. Physiology & Behavior, 112–113(0), 49-55. doi:http://dx.doi.org/10.1016/j.physbeh.2013.02.010

Kolb EM, Rezende EL, Holness L, Radtke A, Lee SK, Obenaus A & Garland TJ. (2013) Mice selectively bred for high voluntary wheel running have larger midbrains: Support for the mosaic model of brain evolution. The Journal of Experimental Biology, 216(Pt 3), 515-523. doi:10.1242/jeb.076000 [doi]

Kong Y, Flick MJ, Kudla AJ, Konieczny SF (1997) Muscle LIM protein promotes myogenesis by enhancing the activity of MyoD. Mol Cell Biol 17: 4750–4760.

Koob GF. (1992) Drugs of abuse: Anatomy, pharmacology and function of reward pathways. Trends in Pharmacological Sciences, 13(5), 177-184.

Kornetsky, C., & Esposito, R. U. (1981). Reward and detection thresholds for brain stimulation: Dissociative effects of cocaine. Brain Research, 209(2), 496-500. doi:0006-8993(81)90177-3 [pii]

Koteja, P, Garland TJ, Sax JK, Swallow JG & Carter PA. (1999) Behaviour of house mice artificially selected for high levels of voluntary wheel running. Animal Behaviour, 58(6), 1307-1318. doi:10.1006/anbe.1999.1270 [doi]

Koteja P, Swallow JG, Carter PA & Garland TJ. (2003) Different effects of intensity and duration of locomotor activity on circadian period. Journal of Biological Rhythms, 18(6), 491-501. doi:10.1177/0748730403256998 [doi]

Kozakowska M, Ciesla M, Stefanska A, Skrzypek K, Was H, et al. (2012) Heme oxygenase-1 inhibits myoblast differentiation by targeting myomirs. Antioxid Redox Signal 16: 113–127.

Krezel W, Ghyselinck N, Samad TA, Dupe V, Kastner P, Borrelli E & Chambon P. (1998) Impaired locomotion and dopamine signaling in retinoid receptor mutant mice. Science (New York, N.Y.), 279(5352), 863-867.

Kroeze WK, Sheffler DJ & Roth BL. (2003) G-protein-coupled receptors at a glance. Journal of Cell Science, 116(24), 4867-4869.

Krzyzosiak A, Szyszka-Niagolov M, Wietrzych M, Gobaille S, Muramatsu S & Krezel W. (2010) Retinoid x receptor gamma control of affective behaviors involves dopaminergic signaling in mice. Neuron, 66(6), 908-920. doi:10.1016/j.neuron.2010.05.004 [doi]

Kuhr WG & Korf J. (1988a) Extracellular lactic acid as an indicator of brain metabolism: Continuous on-line measurement in conscious, freely moving rats with intrastriatal dialysis. Journal of Cerebral Blood Flow and Metabolism : Official Journal of the International Society of Cerebral Blood Flow and Metabolism, 8(1), 130-137. doi:10.1038/jcbfm.1988.17 [doi]

Kuhr WG & Korf J. (1988b) N-methyl-D-aspartate receptor involvement in lactate production following ischemia or convulsion in rats. European Journal of Pharmacology, 155(1-2), 145-149. doi:0014-2999(88)90412-8 [pii]

Labadarios D., McKenzie D.Y., Dickerson J.W. and Parke D.V.. Metabolic abnormalities of tryptophan and nicotinic acid in patients with rheumatoid arthritis. Rheumatol. Rehabil., 17:227-232, 1978.

Lam KS, and Srivastava G. Sex-related differences and thyroid hormone regulation of vasoactive intestinal peptide gene expression in the rat brain and pituitary. Brain research 526: 135-137, 1990.

Lammers CH, D'Souza U, Qin ZH, Lee SH, Yajima S & Mouradian MM. (1999) Regulation of striatal dopamine receptors by estrogen. Synapse (New York, N.Y.), 34(3), 222-227. doi:10.1002/(SICI)1098-2396(19991201)34:3<222::AID-SYN6>3.0.CO;2-J [pii]

Landry M, Levesque D & Di Paolo T. (2002) Estrogenic properties of raloxifene, but not tamoxifen, on D2 and D3 dopamine receptors in the rat forebrain. Neuroendocrinology, 76(4), 214-222. doi:65951 [pii]

Langmead B, Trapnell C, Pop M & Salzberg SL. (2009) Ultrafast and memory efficient alignment of short DNA sequences to the human genome. Genome Biol., 10(3):R25.

Langmead B & Salzberg SL. (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4):357-359.

Le Bechec A, Portales-Casamar E, Vetter G, Moes M, Zindy PJ, et al. (2011) MIR@NT@N: A framework integrating transcription factors, microRNAs and their targets to identify sub-network motifs in a meta-regulation network model. BMC Bioinformatics 12: 67.

Lee, I. C., Ting, T. T., Chen, D. R., Tseng, F. Y., Chen, W. J., & Chen, C. Y. (2015). Peers and social network on alcohol drinking through early adolescence in taiwan. Drug and Alcohol Dependence, 153, 50-58. doi:10.1016/j.drugalcdep.2015.06.010 [doi]

Lee KW, Hong JH, Choi IY, Che Y, Lee JK, Yang SD, Song CW, Kang HS, Lee JH, Noh JS, Shin HS, Han PL. (2002) J. Impaired D2 dopamine receptor function in mice lacking type 5 adenylyl cyclase. J Neurosci. Sep 15;22(18):7931-40.

Lee S, Lee W, Shin J, Han B, Moon S, Cho S, Park T, Kim H, and Han J. Sexually dimorphic gene expression in the chick brain before gonadal differentiation. Poultry science 88: 1003-1015, 2009.

Lee SJ, Reed A, Davies M, Girgenrath S, Goad M, et al. (2005) Regulation of muscle growth by multiple ligands signaling through activin type II receptors. Proc Natl Acad Sci USA 102: 18117–18122.

Lee SJ (2004) Regulation of muscle mass by myostatin. Annu Rev Cell Dev Biol 20: 61–86.

Lee SJ (2010) Extracellular Regulation of Myostatin: A Molecular Rheostat for Muscle Mass. Immunol Endocr Metab Agents Med Chem 10: 183–94.

Leuenberger A. (2006) Endorphins, exercise, and addictions: A review of exercise dependence. Impulse: The Premier Journal for Undergraduate Publications in the Neurosciences,

Levine M & Tijan R. Transcription regulation and animal diversity. Nature, 424:147-151, 2003.

Levran O, Peles E, Randesi M, Correa da Rosa J, Ott J, Rotrosen J . . . Kreek MJ (2015) Synaptic plasticity and signal transduction gene polymorphisms and vulnerability

to drug addictions in populations of european or african ancestry. CNS Neuroscience & Therapeutics, 21(11), 898-904. doi:10.1111/cns.12450 [doi]

Li B & Dewey CN. (2011) RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC bioinformatics 12: 323.

Lim S, Ha J, Choi SW, Kang SG & Shin YC. (2012) Association study on pathological gambling and polymorphisms of dopamine D1, D2, D3, and D4 receptor genes in a korean population. Journal of Gambling Studies / Co-Sponsored by the National Council on Problem Gambling and Institute for the Study of Gambling and Commercial Gaming, 28(3), 481-491. doi:10.1007/s10899-011-9261-1 [doi]

Lim S, Ha J, Choi SW, Kang SG, Shin YC. (2012) Association study on pathological gambling and polymorphisms of dopamine D1, D2, D3, and D4 receptor genes in a Korean population. J Gambl Stud. 28(3):481-91.

Liu J, Merkle FT, Gandhi AV, Gagnon JA, Woods IG, Chiu CN . . . Prober DA. (2015) Evolutionarily conserved regulation of hypocretin neuron specification by Lhx9. Development (Cambridge, England), 142(6), 1113-1124. doi:10.1242/dev.117424 [doi]

Liu XF, Bera TK, Liu LJ, and Pastan I. A primate-specific POTE-actin fusion protein plays a role in apoptosis. Apoptosis 14: 1237-1244, 2009.

Ljungberg T, Apicella P & Schultz W. (1991) Responses of monkey midbrain dopamine neurons during delayed alternation per- formance. Brain Res. 586, 337–341.

Lovmar L, Ahlford A, Jonsson M, and Syvänen A-C. Silhouette scores for assessment of SNP genotype clusters. BMC genomics 6: 35, 2005.

Lupica CR & Riegel AC. (2005) Endocannabinoid release from midbrain dopamine neurons: A potential substrate for cannabinoid receptor antagonist treatment of addiction. Neuropharmacology, 48(8), 1105-1116. doi:S0028-3908(05)00126-7 [pii]

Lutter D, Marr C, Krumsiek J, Lang EW, Theis FJ. (2010) Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. BMC Genomics 11: 224.

MacKenzie MG, Hamilton DL, Pepin M, Patton A, Baar K (2013) Inhibition of myostatin signaling through Notch activation following acute resistance exercise. PLoS ONE 8: e68743.

Majdak P, Bucko PJ, Holloway AL, Bhattacharya TK, DeYoung EK, Kilby CN . . . Rhodes JS. (2014) Behavioral and pharmacological evaluation of a selectively bred mouse model of home cage hyperactivity. Behav Genet, 55, 516-534.

Maldonado R, Robledo P, Chover AJ, Caine SB & Koob GF (1993). D1 dopamine receptors in the nucleus accumbens modulate cocaine self-administration in the rat. Pharmacology, Biochemistry, and Behavior, 45(1), 239-242. doi:0091-3057(93)90112-7

Maldonado R, Saiardi A, Valverde O, Samad TA, Roques BP & Borrelli E. (1997) Absence of opiate rewarding effects in mice lacking dopamine D2 receptors. Nature, 388(6642), 586-589. doi:10.1038/41567 [doi]

Mandal K, Sinea R, Mishra S, Senapati P, and Mohan M. Estimates of genetic parameters for some egg production traits in Japanese quail (Coturnix coturnix japonica). Indian Journal of Animal Health 33: 49-54, 1994.

Marioni JC, Mason CE, Mane SM, Stephens M, and Gilad Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18: 1509-1517, 2008.

Marks AR, McIntyre JO, Duncan TM, Erdjument-Bromage H, Tempst P, et al. (1992) Molecular cloning and characterization of (R)-3-hydroxybutyrate dehydrogenase from human heart. J Biol Chem 267: 15459–63.

Martin A, Ochagavia M, Rabasa L, Miranda J, Fernandez-de-Cossio J & Bringas R. (2010) BisoGenet: A new tool for gene network building, visualization and analysis. BMC Bioinformatics, 11(1), 1-9. doi:10.1186/1471-2105-11-91

Matsakas A, Macharia R, Otto A, Elashry MI, Mouisel E, Romanello V, Sartori R, Amthor H, Sandri M, Narkar V, Patel K (2012) Exercise training attenuates the hypermuscular phenotype and restores skeletal muscle function in the myostatin null mouse. Exp Physiol 97: 125–140.

McCroskery S, Thomas M, Maxwell L, Sharma M, Kambadur R (2003) Myostatin negatively regulates satellite cell activation and self-renewal. J Cell Biol 162: 1135–1147.

McEwen B. (2002) Estrogen actions throughout the brain. Recent Progress in Hormone Research, 57, 357-384.

McKim WA. (2003) Drugs and behavior: An introduction to behavioral pharmacology (5th ed.). New Jersey: Prentice Hall.

McPherron AC & Lee SJ. (1997) Double muscling in cattle due to mutations in the Myostatin gene. Proc. Natl. Acad. Sci. USA, 94: 12457-12461.

McPherron AC, Lawler AM, Lee SJ. (1997) Regulation of skeletal muscle mass in mice by a new TGF-β superfamily member. Nature 234: 83–89.

Mellor AL & Munn DH. (2004) IDO expression by dendritic cells: tolerance and tryptophan catabolism. Nature Rev. Immunol., 4:762-774.

Meussen R, Smolders I, Sarre S et al. (1997) Endurance training effects on neurotransmitter release in rat striatum: an in vivo microdialysis study. Acta Physiol. Scand., 159(4):335-341.

Mills A, Herron K, Bain M, Solomon S, and Faure J. (1994) Eggshell quality in Japanese quail Coturnix Japonica genetically selected for high or low levels of fearfulness. In: Proceedings of the 9th European Poultry Conference, Glasgow, UK1, p. 292-293.

Mills AD, Crawford LL, Domjan M, and Faure JM. (1997) The behavior of the japanese or domestic quail (Coturnix japonica). Neuroscience & Biobehavioral Reviews 21: 261-281.

Minvielle F, Gourichon D, and Moussu C. (2005) Two new plumage mutations in the Japanese quail. BMC genetics 6: 14.

Mirenowicz J & Schultz W. (1994) Importance of unpredictabil- ity for reward responses in primate dopamine neurons. J. Neuro- physiol. 72, 1024–1027.

Miyaki CY, Griffiths R, Orr K, Nahum LA, Pereira SL, and Wajntal A. (1998) Sex Identification of Parrots, Toucans, and. Zoo Biology 17: 415-423.

Mizutani M. (2003) The Japanese quail. Age 80: 90.

Morris CM, Hao QL, Heisterkamp N, Fitzgerald PH and Groffen J. (1991) Localization of the TRK proto-oncogene to human chromosome bands 1q23-1q24. Oncogene, 6(6):1093-1095.

Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B. (2008) Mapping and quantifying mammalian transcriptomes by RNA seq. Nature Methods, 5(7) 621-628.

Mosler S, Relizani K, Mouisel E, Amthor H, Diel P. (2014) Combinatory effects of siRNA-induced myostatin inhibition and exercise on skeletal muscle homeostasis and body composition. Physiol Rep 20;2(3):e00262.

Nadal-Ginard B & Mahdavi V. (1989) Molecular basis of cardiac performance: plasticity of the myocardium generated through protein isoform switches. J Clin Invest 84: 1693–1700.

Nadeau NJ, Minvielle F, Ito Si, Inoue-Murayama M, Gourichon D, Follett SA, Burke T, and Mundy NI. (2008) Characterization of Japanese quail yellow as a genomic deletion upstream of the avian homolog of the mammalian ASIP (agouti) gene. Genetics 178: 777-786.

Nam K, and Ellegren H. The chicken (Gallus gallus) Z chromosome contains at least three nonlinear evolutionary strata. Genetics 180: 1131-1136, 2008.

Nathanson NM. (1987) Molecular properties of the muscarinic acetylcholine receptor. Annual Review of Neuroscience, 10, 195-236. doi:10.1146/annurev.ne.10.030187.001211 [doi]

Nehrenberg DL, Hua K, Estrada-Smith D, Garland TJ & Pomp D. (2009) Voluntary exercise and its effects on body composition depend on genetic selection history. Obesity (Silver Spring, Md.), 17(7), 1402-1409. doi:10.1038/oby.2009.51 [doi]

Nelson DR. (1999) A second CYP26 P450 in humans and zebrafish: CYP26B1. Arch. Biochem. Biophys., 371 (2):345-347.

Nirenberg MJ & Waters C. (2006) Compulsive eating and weight gain related to dopamine agonist use. Mov Disord. 21(4):524–9.

Nixon SE, Gonzalez-Pena D, Lawson MA, McCusker RH, Hernandez AG, O'Connor JC . . . Rodriguez-Zas SL. (2015) Analytical workflow profiling gene expression

in murine macrophages. Journal of Bioinformatics and Computational Biology, 13(2), 1550010. doi:10.1142/S0219720015500109 [doi]

Noble EP. (2003) D2 dopamine receptor gene in psychiatric and neurologic disorders and its phenotypes. American Journal of Medical Genetics.Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics, 116B(1), 103-125. doi:10.1002/ajmg.b.10005 [doi]

Noma T, Fujisawa K, Yamashiro Y, Shinohara M, Nakazawa A, et al. (2001) Structure and expression of human mitochondrial adenylate kinase targeted to the mitochondrial matrix. Biochem J 358: 225–232.

Noonan MA, Bulin SE, Fuller DC & Eisch AJ. (2010) Reduction of adult hippocampal neurogenesis confers vulnerability in an animal model of cocaine addiction. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 30(1), 304-315. doi:10.1523/JNEUROSCI.4256-09.2010 [doi]

O'Connor JC, Lawson MA, Andre C, Briley EM, Szegedi SS, Lestage J, Castanon N, Herkenham M, Dantzer R and Kelley KW. (2009) Induction of IDO by bacille Calmette-Guerin is responsible for development of murine depressive-like behavior. J. Immunol., 182:3202-3212..

Ocaya PA, Elmabsout AA, Olofsson PS, Torma H, Gidlof AC and Sirsjö A. (2010) CYP26B1 plays a major role in the regulation of all-trans-retinoic acid metabolism and signaling in human aortic smooth muscle cells. J. Vasc. Res., 48:23-30.

Oliveros JC. (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html.

Ordway GA & Garry DJ.(2004) Myoglobin: an essential hemoprotein in striated muscle. J Exp Biol 207: 3441–3446.

Oshlack A & Wakefield MJ. (2009) Transcript length bias in RNA-seq data confounds systems biology. Biology Direct, 4:14.

Pajcini KV, Pomerantz JH, Alkan O, Doyonnas R, Blau HM. (2008) Myoblasts and macrophages share molecular components that contribute to cell-cell fusion. J Cell Biol 180: 1005–1019.

Paoletti L, Elena C, Domizi P, and Banchio C. (2011) Role of Phosphatidylcholine during Neuronal differentiation. IUBMB life 63: 714-720.

Park CH, Kang JS, Yoon EH, Shim JW, Suh-Kim H & Lee SH. (2008) Proneural bHLH neurogenin 2 differentially regulates Nurr1-induced dopamine neuron differentiation in rat and mouse neural precursor cells in vitro. FEBS Letters, 582(5), 537-542. doi:10.1016/j.febslet.2008.01.018 [doi]

Patel RK, Jain M (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE 7: e30619.

Penney JB & Young AB (1983) Speculations on the functional anatomy of basal ganglia disorders. Annual Review of Neuroscience, 6, 73-94. doi:10.1146/annurev.ne.06.030183.000445 [doi]

Peoples LL, Gee F, Bibi R & West MO. (1998) Phasic firing time locked to cocaine self-infusion and locomotion: Dissociable firing patterns of single nucleus accumbens neurons in the rat. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 18(18), 7588-7598.

Personius KE, Jayaram A, Krull D, Brown R, Xu T, et al. (2010) Grip force, EDL contractile properties, and voluntary wheel running after postdevelopmental myostatin depletion in mice. J Appl Physiol 109: 886–894.

Phillips GD, Robbins TW & Everitt BJ (1994). Bilateral intra-accumbens self-administration of d-amphetamine: Antagonism with intra-accumbens SCH-23390 and sulpiride. Psychopharmacology, 114(3), 477-485.

Pieau C & Dorizzi M (2004). Oestrogens and temperature-dependent sex determination in reptiles: All is in the gonads. The Journal of Endocrinology, 181(3), 367-377.

Pierce EF, Eastman NW, Tripathi HL, Olson KG, Dewey WL. (1993) Beta-endorphin response to endurance exercise: relationship to exercise dependence. Percept Mot Skills 77:767-770.

Poleszak E & Malec D. (2002) Cocaine-induced hyperactivity is more influenced by adenosine receptor agonists than amphetamine-induced hyperactivity. Polish Journal of Pharmacology, 54(4), 359-366.

Pontone G, Williams JR, Bassett SS, Marsh L. (2006) Clinical features associated with impulse control disorders in Parkinson disease. Neurology. 67(7):1258–61.

Powers SK & Jackson MJ. (2008) Exercise-induced oxidative stress: Cellular mechanisms and impact on muscle force production. Physiol Rev 88:1243–1276.

Ravenscroft G, Colley SM, Walker KR, Clement S, Bringans S, et al. (2008) Expression of cardiac alpha-actin spares extraocular muscles in skeletal muscle alpha-actin diseases–quantification of striated alpha-actins by MRM-mass spectrometry. Neuromuscul Disord 18: 953–958.

Rawat A, Gust KA, Elasri MO, and Perkins EJ. (2010) Quail Genomics: a knowledgebase for Northern bobwhite. BMC bioinformatics 11: S13.

Rhodes JS, Gammie SC & Garland T. (2005) Neurobiology of mice selected for high voluntary wheel-running activity. Integrative and Comparative Biology, 45(3), 438-455. doi:10.1093/icb/45.3.438

Ricklefs R. (1973) Patterns of growth in birds. II. Growth rate and mode of development. Ibis 115: 177-201.

Rios RI, Carneiro I, Arce VM, Devesa J. (2002) Myostatin is an inhibitor of myogenic differentiation. Am J Physiol Cell Physiol 282: C993–C999.

Romo R & Schultz W. (1990) Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self- initiated arm movements. J. Neurophysiol. 63, 592–606.

Rushforth AM, White CC, Anderson P. (1998) Functions of the Caenorhabditis elegans regulatory myosin light chain genes mlc-1 and mlc-2. Genetics 150: 1067–1077.

Salmons S & Vrbova G. (1969) The influence of activity on some contractile characteristics of mammalian fast and slow muscles. J Physiol (Lond.). 201: 535–549.

Samad TA, Krezel W, Chambon P & Borrelli E. (1997) Regulation of dopaminergic pathways by retinoids: Activation of the D2 receptor promoter by members of the retinoic acid receptor-retinoid X receptor family. Proceedings of the National Academy of Sciences of the United States of America, 94(26), 14349-14354.

Samson F, de Jong PJ, Trask BJ, Koza-Taylor P, Speer MC, et al. (1992) Assignment of the human slow skeletal troponin T gene to 19q13.4 using somatic cell hybrids and fluorescence in situ hybridization analysis. Genomics 13: 1374–5.

Savage KJ, McPherron AC (2010) Endurance exercise training in myostatin null mice. Muscle Nerve 42: 355–62.

Schasfoort EM, De Bruin LA & Korf J (1988). Mild stress stimulates rat hippocampal glucose utilization transiently via NMDA receptors, as assessed by lactography. Brain Research, 475(1), 58-63. doi:0006-8993(88)90198-9 [pii]

Schiaffino S, Dyar KA, Ciciliot S, Blaauw B, Sandri M (2013) Mechanisms regulating skeletal muscle growth and atrophy. FEBS J 280: 4294–4314.

Schneeberger D & Raabe T. (2003) Mbt, a Drosophila PAK protein, combines with Cdc42 to regulate photoreceptor cell morphogenesis. Development 130: 427–437.

Schneider AG, Sultan KR, Pette D. (1999) Muscle LIM protein: expressed in slow muscle and induced in fast muscle by enhanced contractile activity. Am J Physiol 276(4 Pt 1):C900-6.

Schroecksnadel K, Winkler C, Duftner C, Wirleitner B, Schirmer M and Fuchs D. (2006) Tryptophan degradation increases with stage in patients with rheumatoid arthritis. Clin. Rheumatol., 25: 334-337.

Schultz W, Apicella P and Ljungberg T. (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J. Neurosci. 13, 900–913.

Self DW, Barnhart WJ, Lehman DA & Nestler EJ. (1996) Opposite modulation of cocaine-seeking behavior by D1- and D2-like dopamine receptor agonists. Science (New York, N.Y.), 271(5255), 1586-1589.

Serao, N. V., Gonzalez-Pena, D., Beever, J. E., Bollero, G. A., Southey, B. R., Faulkner, D. B., & Rodriguez-Zas, S. L. (2013). Bivariate genome-wide association analysis of the growth and intake components of feed efficiency. PloS One, 8(10), e78530. doi:10.1371/journal.pone.0078530 [doi]

Serao NV, Gonzalez-Pena D, Beever JE, Faulkner DB, Southey BR & Rodriguez-Zas SL. (2013) Single nucleotide polymorphisms and haplotypes associated with feed efficiency in beef cattle. BMC Genetics, 14, 94-2156-14-94. doi:10.1186/1471-2156-14-94 [doi]

Short JL, Ledent C, Drago J & Lawrence AJ (2006). Receptor crosstalk: Characterization of mice deficient in dopamine D1 and adenosine A2A receptors.

Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology, 31(3), 525-534. doi:1300852 [pii]

Smith MC, Finger JH, Kadin JA, Richardson JE, Rungwald M (2014) The Gene Expression Database for Mouse development (GXD): Putting developmental expression information at your fingertips. Developmental Dynamics 243:1176-1186.

Soler H, Vinayak P & Quadagno D (2000). Biosocial aspects of domestic violence. Psychoneuroendocrinology, 25(7), 721-739. doi:S0306-4530(00)00022-6 [pii]

Son YJ, Jeong S, Kang BG, Kim SH & Lee SK. (2015) Visualization of e-health research topics and current trends using social network analysis. Telemedicine Journal and E-Health : The Official Journal of the American Telemedicine Association, 21(5), 436-442. doi:10.1089/tmj.2014.0172 [doi]

Springer MS, Stanhope MJ, Madsen O, de Jong WW. (2004) Molecules consolidate the placental mammal tree. Trends Ecol. Evol. 19 (8):430-438.

Sriram S, Subramanian S, Juvvuna PK, McFarlane C, Salerno MS, Kambadur R, Sharma M. (2014) Myostatin induces DNA damage in skeletal muscle of streptozotocin-induced type 1 diabetic mice. J Biol Chem 289: 5784–5798.

Sriram S, Subramanian S, Sathiakumar D, Venkatesh R, Salerno MS, et al. (2011) Modulation of reactive oxygen species in skeletal muscle by myostatin is mediated through NF-κB. Aging Cell 10: 931–948.

Starck CS, Sutherland-Smith AJ (2010) Cytotoxic aggregation and amyloid formation by the myostatin precursor protein. PLoS ONE 5: e9170.

Storey JD. (2002) A direct approach to false discovery rates. J R Statist Soc Ser B 64: 479–498

Storey JD, Taylor JE, Siegmund D (2004) Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: A unified approach. J R Statist Soc Ser B 66: 187–205.

Stumpf MP, Thorne T, de Silva E, Stewart R, An HJ, Lappe M and Wiuf C. (2008) Estimating the size of the human interactome. Proc. Natl. Acad. Sci. USA, 105:6959-6964.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA . . . Mesirov JP. (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences, 102(43), 15545-15550. doi:10.1073/pnas.0506580102

Suh A, Kriegs JO, Brosius J, and Schmitz J. (2011) Retroposon insertions and the chronology of avian sex chromosome evolution. Molecular biology and evolution 28: 2993-2997.

Sussman S, Lisha N & Griffiths M. (2011) Prevalence of the addictions: A problem of the majority or the minority? Evaluation & the Health Professions, 34(1), 3-56. doi:10.1177/0163278710380124 [doi]

Sussman S, Lisha N, Griffiths M: Prevalence of the addictions: a problem of the majority or the minority? Eval Health Prof 2011;34:3–56.

Sutoo D & Akiyaman K. (2003) Regulation of brain function by exercise. Neurobiology of Disease, 13(1), 1-14. doi:S0969996103000305 [pii]

Swallow J, Carter P & Garland TJ. (1998) Artificial selection for increased wheel-running behavior in house mice. Behavior Genetics, 28(3), 227-237. doi:10.1023/A:1021479331779

Takada F, Vander Woude DL, Tong HQ, Thompson TG, Watkins SC, et al. 2001. Myozenin: an alpha-actinin- and gamma- filamin-binding protein of skeletal muscle Z lines. Proc Natl Acad Sci USA 98: 1595–600.

Talmadge R, Acosta W & Garland TJ. (2014) Myosin heavy chain isoform expression in adult and juvenile mini-muscle mice bred for high-voluntary wheel running. Mechanisms of Development, 134(0), 16-30. doi:http://dx.doi.org/10.1016/j.mod.2014.08.004

Teixeira AM, Reckziegel P, Müller L, Pereira RP, Roos DH, Rocha JBT & Bürger ME. (2009) Intense exercise potentiates oxidative stress in striatum of reserpine-treated animals. Pharmacology Biochemistry and Behavior, 92(2), 231-235. doi:http://dx.doi.org/10.1016/j.pbb.2008.11.015

Timson DJ (2003) Fine tuning the myosin motor: the role of the essential light chain in striated muscle myosin. Biochimie 85: 639–645.

Tobon KE, Catuzzi JE, Cote SR, Sonaike A & Kuzhikandathil EV. (2015) Post-transcriptional regulation of dopamine D1 receptor expression in caudate-putamen of cocaine-sensitized mice. The European Journal of Neuroscience, 42(2), 1849-1857. doi:10.1111/ejn.12933 [doi]

Touart L. (2003) Revised draft detailed review paper for avian two-generation toxicity test.

Trapnell C, Pachter L & Salzberg SL. (2009) TopHat: discovering splice junctions with RNA-seq. Bioinformatics, 25(9):1105-1111.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc., 7(3): 562-78.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L. (2010) Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnol., 28(5):511-515.

Tsujino N & Sakurai T. (2009) Orexin/hypocretin: A neuropeptide at the interface of sleep, energy homeostasis, and reward system. Pharmacological Reviews, 61(2), 162-176. doi:10.1124/pr.109.001321 [doi]

Tuch BB, Li H & Johnson AD. (2008) Evolution of eukaryotic transcription circuits. Science, 319 (5871):1797-1799.

Udan RS, Culver JC & Dickinson ME. (2013) Understanding vascular development. Wiley Interdiscip Rev Dev Biol 2: 327–346.

Uyttenhove C, Pilotte L, Theate I, Stroobant V, Colau D, Parmentier N, Boon T and Van den Eynde BJ. (2003) Evidence for a tumoral immune resistance mechanism based on tryptophan degradation by indoleamine 2,3-dioxygenase. Nature Med., 9:1269-1274,

Van Verk MC, Hickman R, Pieterse CM, Van Wees SC (2013) RNA-Seq: revelation of the messengers. Trends Plant Sci 18: 175-179. doi:10.1016/j.tplants.2013.02.001.

Varga J, Yufit T & Brown RR. (1995) Inhibition of collagenase and stromelysin gene expression by interferon-gamma in human dermal fibroblasts is mediated in part via induction of tryptophan degradation. J. Clin. Invest., 96:475-481.

Vivier E, Tomasello E, Baratin M, Walzer T & Ugolini S. (2008) Functions of natural killer cells. Nature Immunology, 9(5), 503-510. doi:10.1038/ni1582 [doi]

Voon V, Hassan K, Zurowski M, de Souza M, Thomsen T, Fox S, et al. Prevalence of repetitive and reward-seeking behaviors in Parkinson disease. Neurology 2006;67(7):1254–7.

Wahlberg P, Strömstedt L, Tordoir X, Foglio M, Heath S, Lechner D, Hellström AR, Tixier-Boichard M, Lathrop M, and Gut IG. A high-resolution linkage map for the Z chromosome in chicken reveals hot spots for recombination. Cytogenetic and genome research 117: 22-29, 2007.

Wang GJ, Volkow ND, Fowler JS, Franceschi D, Logan J, Pappas NR . . . Netusil N. (2000) PET studies of the effects of aerobic exercise on human striatal dopamine release. Journal of Nuclear Medicine : Official Publication, Society of Nuclear Medicine, 41(8), 1352-1356.

Wang J, Cui W, Wei J, Sun D, Gutala R, Gu J & Li MD. (2011) Genome-wide expression analysis reveals diverse effects of acute nicotine exposure on neuronal function-related genes and pathways. Frontiers in Psychiatry, 2, 5. doi:10.3389/fpsyt.2011.00005

Wang Z, Gerstein M & Snyder M. (2009) RNA-seq: A revolutionary tool for transcriptomics. Nature Reviews.Genetics, 10(1), 57-63. doi:10.1038/nrg2484 [doi]

Wang D, Saga Y, Mizukami H, Sato N, Nonaka H. et al. (2012) Indoleamine-2,3-dioxygenase, an immunosuppressive enzyme that inhibits natural killer cell function, as a useful target for ovarian cancer therapy. Int. J. Oncol., 40:929-934.

Wang SC, Myers S, Dooms C, Capon R, Muscat GE (2010) An ERRbeta/gamma agonist modulates GRalpha expression, and glucocorticoid responsive gene expression in skeletal muscle cells. Mol Cell Endocrinol 315: 146–152.

Wang Z, Gerstein M, Snyder M (2009) RNA-seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57–63.

Waters RP, Pringle RB, Forster GL, Renner KJ, Malisch JL, Garland TJ & Swallow JG. (2013).= Selection for increased voluntary wheel-running affects behavior and

brain monoamines in mice. Brain Research, 1508(0), 9-22. doi:http://dx.doi.org/10.1016/j.brainres.2013.01.033

Wess J. (1996) Molecular biology of muscarinic acetylcholine receptors. Critical Reviews in Neurobiology, 10(1), 69-99.

Wilson W, Anderson B, and Siopes T. (1971) Importation of wild strain Japanese quail (wild coturnix) offers new game bird possibility. California Agriculture 25: 5-6.

Wise RA. (1996) Neurobiology of addiction. Current Opinion in Neurobiology, 6(2), 243-251. doi:S0959-4388(96)80079-1 [pii]

Wise RA. (2002) Brain reward circuitry: insights from unsensed incentives. Neuron 36, this issue, 229–240.

Wise RA and Hoffman DC. (1992). Localization of drug reward mechanisms by intracranial injections. Synapse 10, 247–263.

Wittenberg JB (1970). Myoglobin-facilitated oxygen diffusion: role of myoglobin in oxygen entry into muscle. Physiol Rev 50: 559–636.

Wolfman NM, McPherron AC, Pappano WN, Davies MV, Song K, et al. (2003) Activation of latent myostatin by the BMP-1/tolloid family of metalloproteinases. Proc Natl Acad Sci USA 100: 15842–15846.

Wone B, Donovan ER, Hayes JP. (2011) Metabolomics of aerobic metabolism in mice selected for increased maximal metabolic rate. Comp Biochem Physiol D 6: 399–405.

Xiong Y, Chen X, Chen Z, Wang X, Shi S, Wang X, Zhang J, and He X. (2010) RNA sequencing shows no dosage compensation of the active X-chromosome. Nature genetics 42: 1043-1047, 2010.

Yokoyama K, Tezuka T, Kotani M, Nakazawa T, Hoshina N, Shimoda Y, Kakuta S, Sudo K, Watanabe K, and Iwakura Y. NYAP: a phosphoprotein family that links PI3K to WAVE1 signalling in neurons. The EMBO journal 30: 4739-4754, 2011.

Yu QT, Ifegwu J, Marian AJ, Mares A Jr, Hill R, Perryman MB, Bachinski LL, Roberts R. (1993) Hypertrophic cardiomyopathy mutation is expressed in mRNA of skeletal as well as cardiac muscle. Circulation. 87: 406–412.

Zayed A & Robinson GE. (2012)Understanding the relationship between brain gene expression and social behavior: Lessons from the Honey Bee. Ann. Rev. Genet., 46: 591-615.

Zhang SZ, Xu Y, Xie HQ, Li XQ, Wei WQ, et al. (2009) The possible role of Myosin light chain in myoblast proliferation. Biol Res 42: 121–132.

Zhao B, Li EJ, Wall RJ, et al. (2009) Coordinated patterns of gene expressions for adult muscle build-up in transgenic mice expressing myostatin propeptide. BMC Genomics 10: 305